



**Perception et production des voyelles orales du français
par des futures enseignantes tchèques de Français
Langue Etrangère (FLE)**

Nikola Paillereau

► **To cite this version:**

Nikola Paillereau. Perception et production des voyelles orales du français par des futures enseignantes tchèques de Français Langue Etrangère (FLE). Linguistique. Université Sorbonne Paris Cité, 2015. Français. NNT : 2015USPCA004 . tel-01162192

HAL Id: tel-01162192

<https://theses.hal.science/tel-01162192>

Submitted on 9 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE SORBONNE NOUVELLE - PARIS 3

École Doctorale 268 « Langage et Langues : Description,
théorisation, transmission »

UFR de Linguistique et Phonétique Générales et Appliquées
(ILPGA)

Thèse de doctorat en Phonétique
pour obtenir le grade de
Docteur de l'Université Sorbonne Nouvelle – Paris 3

Nikola MAUROVÁ PAILLEREAU

**PERCEPTION ET PRODUCTION DES
VOYELLES ORALES DU FRANÇAIS PAR DES
FUTURES ENSEIGNANTES TCHEQUES DE
FRANÇAIS LANGUE ETRANGERE (FLE)**

Thèse dirigée par
Madame Jacqueline Vaissière

Soutenue le 12 janvier 2015

Jury :

Mme Martine ADDA-DECKER, directeur de recherches
M. Jean-Yves DOMMERGUES, professeur émérite (rapporteur)
M. Tomáš DUBĚDA, professeur adjoint
M. Pierre HALLÉ, directeur de recherches
M. Rudolph SOCK, professeur
Mme Jacqueline VAISSIÈRE, professeur (directeur de thèse)
M. François WIOLAND, professeur émérite (rapporteur)

Résumé

Cette étude acoustico-perceptive concerne les limites de la perception et de la production des voyelles orales du français [i, e, ε, a, u, o, ɔ, y, ø, œ], en isolation et en contextes consonantiques divers, chez dix tchécophones, futures enseignantes de Français Langue Étrangère (FLE). Les résultats montrent que (1) La maîtrise phonétique des voyelles dépend de leurs graphies et de l'entourage consonantique. (2) Les voyelles fermées [i, y, u] et le [a] sont globalement maîtrisées avec authenticité. (3) Les capacités de perception des contrastes entre les voyelles moyennes e/ε, ø/œ et o/ɔ ainsi que leur production sont limitées. Ces résultats ne sont que partiellement en accord avec les prédictions établies à partir du Speech Learning Model (SLM) de Flege (1995), basé sur la notion de *similarité* phonétique qui existe entre la langue maternelle (LM) et la langue étrangère (LE).

Mots clés : *phonétique, acoustique, formants, coarticulation, perception, prononciation, voyelles, français, tchèque, Français Langue Étrangère (FLE), Speech Learning Model (SLM)*

Abstract

This acoustic-perceptual study concerns the limits of perception and production of French oral vowels [i, e, ε, a, u, o, ɔ, y, ø, œ], in isolation and in different consonantal contexts, in ten pre-service Czech teachers of French as a Foreign Language (FFL). The results show that (1) Phonetic proficiency in vowels depends on their spellings and consonantal context. (2) Vowels [i, y, u] and [a] are generally mastered with authenticity. (3) The ability to hear contrasts between the vowels e/ε, ø/œ and o/ɔ and pronounce them is limited. These results are only partially consistent with the predictions established in the Speech Learning Model (SLM) by Flege (1995), based on the notion of phonetic *similarity* between the mother tongue (MT) and the foreign language (FL).

Keywords: *phonetics, acoustics, formants, coarticulation, perception, pronunciation, vowels, French, Czech, FFL (French as a Foreign Language), SLM (Speech Learning Model)*

A maman, Benjamin et Adi

Remerciements

Il était une fois une petite fille, Hanka, née en Tchécoslovaquie dans une famille pauvre. Hanka se portait bien, très bien même, et ses parents étaient heureux et fiers qu'elle soit si ronde. Hanka, en revanche, ne supportait pas son embonpoint et elle décida, à l'âge de douze ans, que puisqu'elle était grosse, il fallait qu'elle soit intéressante. Elle se mit à apprendre des langues étrangères : d'abord l'allemand, puis le français, ensuite l'anglais, le russe, le japonais, l'italien et l'espagnol. Bien sûr, elle ne prononce pas toutes ces langues comme les natifs mais cela n'a rien d'étonnant, puisqu'elle a appris la plupart de ces langues en autodidacte, avec des livres. A cette époque-là, dans son pays, il était difficile de trouver des professeurs d'une langue autre que le russe... Sept ans après avoir commencé à apprendre le français, elle entendit parler, pour la première fois, un VRAI Français et elle fut émue en réalisant que le français était une langue vivante !

Hanka est ma mère, et quand je suis née, nous avons quitté mon père, car il était davantage aventurier que père de famille. Ma mère m'a élevée seule et comme elle n'avait pas assez d'argent pour me payer une nounou, elle m'embarquait à son travail, dans l'école des langues où elle enseignait. Depuis ma petite enfance, j'ai donc eu droit à de gros bains linguistiques et je n'ai pu faire autrement que de m'intéresser, à mon tour, aux langues étrangères. Dans un premier temps, l'anglais avait ma préférence. Mais après avoir rencontré Benjamin, mon futur mari, je me suis plongée dans le français et j'y nage toujours. Cette alliance franco-tchèque a donné naissance à une magnifique petite fille Adi et à cette thèse que vous allez lire. La thèse est une longue histoire, combinant des périodes d'enthousiasme et de déclin, comme souvent j'imagine. Mais malgré tout, il s'agit d'une belle histoire, qui m'a appris, en dehors bien sûr des savoirs en phonétique, à développer persévérance, discipline et modestie.

Je voudrais remercier toute ma famille de m'avoir soutenue. Merci à mon mari d'avoir passé ses soirées et ses week-ends à écrire le programme VisuVo qui m'a énormément facilité la visualisation des résultats acoustiques. Merci à ma mère qui, elle, n'a pas terminé sa thèse et qui aurait mérité son PhD. Merci aussi à ma fille avec qui j'ai passé moins de temps que j'aurais souhaité et qui est devenue le moteur turbo de ce travail. Rassure-toi, ma fille, tout le monde ne travaille pas à la bibliothèque. J'adresse également une pensée à mon père que je ne vois plus mais qui était là au début de la thèse et qui m'a soutenue tout le long de mon Master.

Un grand merci à ma directrice, Mme Vaissière qui m'a motivée à me lancer dans cette aventure. Je voudrais notamment saluer sa précision et sa volonté d'obtenir le maximum de ses étudiants et pour ses étudiants.

Je remercie ensuite tous les locuteurs et auditeurs qui m'ont permis de réaliser ce travail. Merci aux collègues de la fac, Naomi, Lucille, Christine, Céline, Marine, Marie, Agathe, Angélique, Laurence, Nico... Merci aussi à la famille et aux amis qui ont participé aux expériences, Marino, Jean-Yves, Kubo, Xavier, Estelle, Marie-Astrid, Faustyne, Hanko et Lucko.

Je tiens à remercier les futures enseignantes de FLE de LM tchèque, sans qui cette thèse n'aurait pas été possible, Bety, Dany, les deux Eva, Hanka, Lenka, Kristyna, Misa, Mirka et Zuzka. Je vous suis reconnaissante. Merci à Sylva de m'avoir donné ces contacts.

Enfin, je remercie chaleureusement mes relecteurs, Florentina, Simon, Takeki et Matthieu.

Le fait d'avoir terminé cette thèse clôt également une partie de ma vie. Mais après chaque fin, il y a toujours un début et comme l'a dit Kundera (1987), « *il n'est rien de plus beau que l'instant qui précède le voyage, l'instant où l'horizon de demain vient nous rendre visite et nous dire ses promesses.* ».

Merci et bonne lecture.

Table des matières

LISTE DES SIGLES ET ABREVIATIONS	13
AVANT-PROPOS	15
PRESENTATION DE L'ETUDE	17
<i>Structure de l'étude</i>	<i>18</i>
PREMIERE PARTIE : INTRODUCTION.....	23
1. CARACTERISATION DES VOYELLES SUR LES PLANS ARTICULATOIRE, ACOUSTIQUE ET PERCEPTIF	25
1.1 <i>Théorie acoustique de la production de la parole.....</i>	<i>25</i>
1.2 <i>Coarticulation</i>	<i>31</i>
1.3 <i>Perception du timbre vocalique</i>	<i>44</i>
2. LE TCHEQUE ET LE FRANÇAIS	49
2.1 <i>Codes grapho-phoniques</i>	<i>50</i>
2.2 <i>Systèmes segmentaux</i>	<i>50</i>
2.3 <i>Systèmes supra-segmentaux.....</i>	<i>63</i>
3. APPRENTISSAGE PHONETIQUE	67
3.1 <i>Objets de l'apprentissage phonétique.....</i>	<i>68</i>
3.2 <i>Difficulté d'apprentissage des voyelles.....</i>	<i>71</i>
3.3 <i>Etudes antérieures sur l'apprentissage phonétique des voyelles du français</i>	<i>74</i>
3.4 <i>Théories de l'apprentissage phonétique.....</i>	<i>79</i>
4. ENSEIGNEMENT DE LA PHONETIQUE.....	87
4.1 <i>Quelques outils de l'enseignement phonétique.....</i>	<i>88</i>
4.2 <i>Norme phonétique.....</i>	<i>91</i>
4.3 <i>Méthodes d'enseignement des langues étrangères.....</i>	<i>92</i>
4.4 <i>Evaluation du niveau phonétique.....</i>	<i>95</i>
5. ACCENT ETRANGER	97
5.1 <i>Age au début de l'apprentissage</i>	<i>98</i>
5.2 <i>Qualité et quantité d'input.....</i>	<i>99</i>
5.3 <i>Utilisation relative de la L2 par rapport à la L1</i>	<i>100</i>
5.4 <i>Nombre d'années de résidence dans le pays où est parlée la L2</i>	<i>100</i>
5.5 <i>Genre</i>	<i>100</i>
5.6 <i>Types d'apprentissage</i>	<i>101</i>
5.7 <i>Facteurs personnels.....</i>	<i>102</i>
5.8 <i>Perception par le natif.....</i>	<i>102</i>
5.9 <i>Propriétés phonético-phonologiques des sons.....</i>	<i>102</i>
5.10 <i>Conséquences de l'accent étranger</i>	<i>103</i>
6. METHODOLOGIE.....	105
6.1 <i>Locutrices.....</i>	<i>105</i>
6.2 <i>Corpus.....</i>	<i>108</i>
6.3 <i>Passation de l'expérience en production</i>	<i>110</i>
6.4 <i>Base de données.....</i>	<i>112</i>
6.5 <i>Traitement des données.....</i>	<i>114</i>
6.6 <i>Visualisation des résultats acoustiques.....</i>	<i>115</i>
6.7 <i>Interprétation des résultats acoustiques.....</i>	<i>118</i>
6.8 <i>Tests de perception et leur interprétation.....</i>	<i>120</i>
6.9 <i>Le choix du modèle d'apprentissage phonétique.....</i>	<i>122</i>
DEUXIEME PARTIE : COMPARAISON DES VOYELLES ORALES DU FRANÇAIS NON-MERIDIONAL ET DU TCHEQUE DE BOHEME	123
7. SYMBOLES VOCALIQUES (API)	125
7.1 <i>En français.....</i>	<i>126</i>
7.2 <i>En tchèque.....</i>	<i>126</i>
8. PROPRIETES ACOUSTIQUES DES VOYELLES ISOLEES.....	129

8.1	<i>Valeurs formantiques des cibles</i>	129
8.2	<i>Focalisation formantique</i>	134
8.3	<i>Contrastes vocaliques</i>	136
8.4	<i>Similarité acoustique des voyelles isolées du français et du tchèque</i>	148
9.	PROPRIETES ACOUSTIQUES DES VOYELLES EN CONTEXTES PHONETIQUES P, T, K, R/H : EFFETS DE LA COARTICULATION	153
9.1	<i>Prédictions à partir d'un modèle articulatoire (modèle Maeda)</i>	154
9.2	<i>Coarticulation en français : Données réelles</i>	155
9.3	<i>Coarticulation en tchèque : Données réelles</i>	178
9.4	<i>Différences entre les effets de la coarticulation en français et en tchèque</i>	192
9.5	<i>Similarité acoustique des voyelles du français et du tchèque en contextes symétriques p, t, k</i>	196
10.	PROPRIETES ACOUSTIQUES DES VOYELLES INITIALES, MEDIANES ET FINALES DES LOGATOMES : EFFETS DE LA POSITION PROSODIQUE	199
10.1	<i>Réduction due à la position prosodique en français</i>	201
10.2	<i>Réduction due à la position prosodique en tchèque</i>	208
10.3	<i>Réduction due à la position prosodique en français et en tchèque : Conclusion partielle</i>	215
11.	PROPRIETES PERCEPTIVES DES VOYELLES : PERCEPTION INTERLANGUE DES VOYELLES DU FRANÇAIS PAR DES TCHEQUES MONOLINGUES	217
11.1	<i>Voyelles isolées du français</i>	217
11.2	<i>Voyelles du français en contextes p, t, k</i>	221
	CONCLUSION PARTIELLE SUR LA SIMILARITE DES VOYELLES DU FRANÇAIS NON-MERIDIONAL ET DU TCHEQUE DE BOHEME	231
TROISIEME PARTIE : VOYELLES ORALES DU FRANÇAIS DES TCHECOPHONES PAR RAPPORT A LA « REFERENCE »		235
12.	VOYELLES ORALES DU FRANÇAIS CHEZ DES FUTURES ENSEIGNANTES DE FLE DE LM TCHEQUE	237
12.1	<i>Voyelles nouvelles</i>	242
12.2	<i>Voyelles similaires</i>	246
12.3	<i>Voyelles orales du français en contexte uvulaire (analyse formantique)</i>	267
12.4	<i>Voyelles orales du français non maîtrisées par les futures enseignantes tchèques</i>	269
12.5	<i>Conclusion partielle</i>	273
13.	VOYELLES MOYENNES DU FRANÇAIS CHEZ DES FRANÇAIS NATIFS NON-MERIDIONAUX	275
13.1	<i>Tests de perception</i>	277
13.2	<i>Production de contrastes entre e/ɛ, o/ɔ, ø/œ</i>	295
13.3	<i>Voyelles moyennes chez des Français natifs : Conclusion partielle</i>	307
14.	VOYELLES MOYENNES DU FRANÇAIS CHEZ DES FUTURES ENSEIGNANTES DE FLE DE LM TCHEQUE	309
14.1	<i>Perception: tâche de la singularité</i>	309
14.2	<i>Production: analyse formantique</i>	310
14.3	<i>Voyelles moyennes chez des futures enseignantes tchèques : Conclusion partielle</i>	340
CONCLUSION FINALE ET DISCUSSIONS		341
BIBLIOGRAPHIE		353
LISTE DES TABLEAUX		369
LISTE DES FIGURES		378

Liste des sigles et abréviations

F0 : fréquence fondamentale
F1, F2, F3, F3 : premier, deuxième, troisième, quatrième formants
Hz : Hertz
ms : milliseconde(s)
ET : écart(s) type(s)
C : consonne
V : voyelle
iso : isolée/ isolation
VOT : *Voice Onset Time* : temps d'établissement de voisement
API : Alphabet Phonétique International
tch : tchèque
fr : français
L1/ LM : langue première/ langue maternelle
L2/ LE : langue seconde/ langue étrangère
FLE : Français Langue Étrangère
SLM : *Speech Learning Model* : Modèle d'Apprentissage de la Parole de Flege (1995)
PAM : *Perceptual Assimilation Model* : Modèle d'Assimilation Perceptive de Best *et al.* (1995)
NLM : Native Language Magnet : théorie de l'Aimant perceptif de la Langue Maternelle de Kuhl and Iverson (1995)

Avant-propos

L'admiration que la Bohême éprouve pour la France remonte à l'époque du Moyen-Age. Ceux qui connaissent Prague, la capitale de la République tchèque, savent que de nombreux monuments historiques portent le nom de Charles IV (1316-1378), le roi tchèque et l'empereur romain germanique qui fut élevé à la cour française. A son retour à Prague en 1483, Charles IV fonda, selon le modèle de la Sorbonne, la première université de l'Europe centrale, nommée par la suite l'Université Charles (Hoensch, 1995). Plus tard, sous l'empire des Habsbourg, les relations avec la France s'affaiblirent et ne furent renouées qu'au 19^e siècle. La francophilie des Tchèques se traduisit en 1886 par la fondation à Prague de la première Alliance Française de l'Europe centrale, puis en 1920 par celle de l'Institut français de Prague par Ernest Denis, historien français. De nombreux artistes tchèques partirent pour vivre et créer en France, comme Alphonse Mucha (1860-1939), connu par les Parisiens pour ses peintures qui décorent les entrées de métro, ou Milan Kundera, (né en 1929) l'écrivain d'origine tchèque. Les institutions françaises furent fermées durant l'occupation allemande (1939-1945) et le régime communiste (1948-1989) mais furent rouvertes et l'enseignement du français reprit dès 1989.

Le français est actuellement enseigné à l'Institut Français de Prague (www.ifp.cz) ainsi que dans les six Alliances françaises de la République tchèque (www.alliancefrancaise.cz) et également dans de nombreuses écoles de langues et des lycées. Les Tchèques disent aimer la sonorité du français parlé (« ça fait chic ») ce qui est une des raisons du choix du français dans les collèges et lycées en République tchèque où l'allemand et l'espagnol sont également proposés comme deuxième LE, après l'anglais (Rakova, 2011). Mais est-il possible aux adolescents tchèques qui commencent l'apprentissage de FLE à l'âge de onze ans, de parler le français comme des natifs, c'est-à-dire sans « accent étranger » ?

Afin d'assurer une transmission du français oral le plus authentique possible, les enseignants, qui sont le seul modèle en milieu exolingue (outre les supports audio et vidéo présentant les différentes variétés du français) devraient quasiment parler comme des natifs (Leather, 1983) pour donner aux apprenants des bons exemplaires des sons et de la prosodie à maîtriser. Ils doivent en même temps posséder de nombreuses connaissances théoriques en phonétique (Delattre, 1960; Wachs, 2011) et avoir un modèle didactique. En effet, recourir à un modèle a l'avantage de présenter aux apprenants l'usage réel de la langue de manière simplifiée (Wioland, 2005).

Les manuels de FLE du marché actuel illustrent souvent la production des sons du français de façon classique, c'est-à-dire par des schémas articulatoires. La position de la langue, des lèvres et des mâchoires est relativement facile à préciser dans la description de l'articulation des consonnes, qui sont produites avec une constriction forte, donc localisable dans le conduit vocal (Vaissière, 2007, 2009). En revanche, les schémas articulatoires illustrant l'articulation des voyelles, ayant une constriction moins précise, sont plus approximatifs et ils ne reflètent pas les possibilités de compensations articulatoires dans la production d'une même voyelle. Ainsi, dans l'apprentissage des voyelles françaises, dites difficiles à reproduire authentiquement (Blake, 1972; Racine *et al.*, 2012) et dont le timbre a une place dominante au sein de la syllabe (Delattre, 1953), l'apprenant est invité à viser les cibles acoustiques (Borrell, 1993). Les voyelles sont souvent décrites en termes de la fréquence des deux premiers formants (F1 – F2), ce qui est insuffisant pour en reproduire le timbre exact. Les propriétés nécessaires et suffisantes ont été décrites pour le français par Vaissière (2011), à l'aide de la fréquence des formants supérieurs pour les voyelles antérieures, et de la distance entre les formants : le regroupement de certains formants pour certaines voyelles est essentiel en français pour créer un pic spectral dans une zone bien précise du spectre des fréquences.

Présentation de l'étude

Cette étude acoustico-perceptive concerne les limites de la perception et de la production des voyelles orales du français [i, e, ε, a, u, o, ɔ, y, ø, œ], en isolation et en contextes consonantiques divers, chez dix tchécophones, futures enseignantes de Français Langue Étrangère (FLE). Notre intérêt porte volontairement sur les professionnels car ils sont censés posséder une prononciation quasi-native de la langue enseignée (Leather, 1983).

Pour prédire les limites que nos futures enseignantes peuvent rencontrer dans la maîtrise des voyelles françaises, nous utilisons le Speech Learning Model (SLM) de Flege (1995) qui s'appuie sur la notion de similarité phonétique interlangue. Selon ce modèle, les segments étrangers *nouveaux* et *identiques* peuvent être maîtrisés par des apprenants avancés avec une plus grande précision que les segments étrangers *similaires* aux segments maternels.

La similarité entre les voyelles orales françaises [i, e, ε, a, u, o, ɔ, y, ø, œ] et les voyelles monophthongues tchèques [a, aː, ε, εː, ɪ, iː, o, oː, u, uː] est estimée sur trois plans:

- (1) « Symbolique ». Nous comparons alors les symboles de l'Alphabet Phonétique International (API) utilisés pour transcrire généralement le timbre des voyelles des deux langues. En considérant tous les symboles utilisés dans la littérature pour décrire le système vocalique tchèque, les voyelles du français [y, ø, œ] sont *nouvelles* alors que [a, i, u, e, ε, o, ɔ] sont *similaires* ou *identiques* aux voyelles du tchèque.
- (2) Acoustique. Sur la base de la fréquence de deux à quatre premiers formants, nous étudions la distance acoustique entre les voyelles du français et les voyelles du tchèque les plus proches. Avec un corpus comportant des voyelles en isolation et dans des logatomes CVCVCVC, nous prenons en compte l'effet du contexte phonétique (nul et labial, coronal et palato-vélaire) et de la position prosodique dans CVCVCVC (initiale, finale et médiane) sur les patrons formantiques. Le résultat montre que les formants vocaliques du français et du tchèque changent peu suivant la position dans le logatome mais les deux langues diffèrent par l'ampleur des variations acoustiques dues à la coarticulation. Alors que les voyelles antérieures françaises produites en contexte varient par rapport à la cible dans une moindre mesure que les voyelles antérieures tchèques, les voyelles postérieures sont plus fortement antériorisées en français qu'en tchèque, notamment en contexte coronal. Ainsi, la similarité acoustique des voyelles françaises et tchèques dépend de leur environnement consonantique.
- (3) Perceptif. Par le biais d'un test d'identification interlangue, nous estimons la distance perceptive entre les voyelles des deux langues : dix auditeurs tchèques monolingues ont identifié les voyelles françaises produites isolément et en contexte labial (pVp), coronal (tVt) et palato-vélaire (kVk) en termes de catégories tchèques.

L'ensemble des résultats obtenus montre que : (a) Il n'existe pas de voyelles françaises *identiques* aux voyelles tchèques (il est possible toutefois de qualifier certaines voyelles de « hautement similaires » pour lesquelles Bohn and Flege (1990) prévoient également une maîtrise authentique), (b) Toutes les voyelles françaises sont, au moins dans quelques contextes phonétiques identiques, *similaires* aux voyelles tchèques, (c) Seules les deux voyelles labiales du français [y] et [ø], selon l'environnement phonétique, peuvent être considérées comme *nouvelles* pour les Tchèques. Ce résultat montre la complexité des prédictions concernant l'apprentissage phonétique établies à partir de la similarité phonétique interlangue.

Afin de vérifier l'hypothèse de SLM et d'estimer la maîtrise en production, nous avons étudié les patrons formantiques des voyelles du français produites par nos futures enseignantes isolément et en syllabe finale pVp, tVt et kVk. Nous avons également étudié le degré de précision des voyelles produites en syllabe finale RVR. Pour vérifier perceptivement l'interprétation des résultats acoustiques, les voyelles isolées ont ensuite été identifiées et évaluées par dix Français natifs. La

maîtrise en perception de nos futures enseignantes a été estimée à partir d'identifications des voyelles isolées produites par une Française parisienne.

Le résultat n'est qu'en partie en concordance avec les prédictions de SLM. En effet, dans les contextes où les voyelles du français sont *hautement similaires* à une voyelle tchèque, leur maîtrise par les tchécoslovaques est généralement authentique. À l'inverse, plus la voyelle du français est acoustiquement éloignée du correspondant tchèque le plus proche, moins sa cible est atteinte par les futures enseignantes, comme prévu par le SLM, à l'exception du [i]. Notons cependant l'immense influence de la graphie qui prime en général sur la similarité phonétique. Enfin, la voyelle *nouvelle* [ø] n'est pas maîtrisée avec authenticité, ce qui est contre les attentes du SLM. Ainsi, les voyelles fermées [i, y, u] et la voyelle ouverte [a] sont globalement maîtrisées de manière native alors que les capacités de perception et production des contrastes entre les voyelles moyennes e/ɛ, ø/œ et o/ɔ sont limitées.

L'étude approfondie de la perception et de la production des voyelles moyennes par des francophones natifs non-méridionaux a mis en évidence que les voyelles moyennes sont réalisées au travers de six qualités bien distinctes et que leur distribution obéit en français standard aux règles orthoépiques. Ainsi la voyelle /e/ n'apparaît jamais en syllabe finale fermée alors que /ɔ/ et /œ/ n'apparaissent pas en syllabe finale ouverte. En revanche, quoiqu'il y ait une tendance générale chez nos francophones à distribuer les voyelles moyennes selon la loi de position, les deux couples de voyelles ø/œ et o/ɔ peuvent contraster en syllabe finale fermée (comme dans *pomme/paume* ou *veule* (adj.)/ *veulent*) et le couple e/ɛ en syllabe finale ouverte (comme dans *taie/ thé*). Par conséquent, malgré le faible rendement phonologique des contrastes, les paires minimales où les oppositions sont maintenues par tous les natifs doivent être présentées et maîtrisées par les futures enseignantes de FLE.

L'étude qualitative de la perception et production des voyelles moyennes par nos futures enseignantes séparément a mis en évidence que : (a) Les cibles [ø] et [ɛ] sont produites avec authenticité par la plupart des locutrices, (b) Les cibles [e] et [o] sont produites avec authenticité par une seule locutrice (ce qui reflète la possibilité de surmonter l'influence de la graphie). (c) Le [œ] isolé est maîtrisé par une locutrice, sans authenticité. (d) Le [ɔ] isolé n'est pas maîtrisé. Notons que les résultats en (c) et (d) ne doivent pas être considérés comme un échec dans l'apprentissage phonétique du français car la production des [ɔ] et [œ] isolés enfreint les règles orthoépiques du français standard.

Ensuite, quoique les cibles ne soient pas toujours atteintes ou maîtrisées, certaines locutrices possèdent deux catégories mentales différentes pour les voyelles en contraste, du moins dans certains contextes phonétiques facilitants. Cependant les différences acoustiques entre e/ɛ, ø/œ et o/ɔ produites par les tchécoslovaques risquent de ne pas être phonétiquement pertinentes pour un auditeur français. En effet, elles sont soit moindres que chez les natifs, soit concernent un indice acoustique différent. Ainsi, pour observer le progrès dans la réalisation des contrastes non perçus, il est indispensable de visualiser les résultats acoustiques. Dans ce but, nous avons créé le logiciel VisuVo (Visualisation des Voyelles) qui permet de générer en temps réel et de manière interactive, à partir de grandes bases de données, des graphiques illustrant les valeurs formantiques cibles de même que leur évolution en fonction de la position prosodique et du contexte phonétique.

Enfin, le lien entre perception et production est complexe car les futures enseignantes de FLE produisent parfois des voyelles ou contrastes vocaliques du français avec authenticité tout en les identifiant ou les discriminant avec difficulté et vice versa.

En accord avec le SLM, cette thèse a ainsi montré que la coarticulation affecte différemment la similarité phonétique des voyelles françaises et tchèques. Par conséquent, la maîtrise des voyelles françaises par des tchécoslovaques dépend de leurs contextes phonétiques mais aussi de leurs graphies, facteur négligé par Flege (1995).

Structure de l'étude

La structure de la thèse, illustrant la répartition des chapitres, se trouve à la Figure 1 (page 20).

La première partie de la thèse (chapitres 1 - 6) introduit les notions utilisées dans cette étude et constitue une revue de la littérature sur les travaux antérieurs. Cette partie présente également la méthodologie du travail.

Le **chapitre un** définit les voyelles sur les plans articulatoire, acoustique et perceptif. En (1.1), nous rappelons les bases de la théorie acoustique de la parole. Nous définissons ensuite la coarticulation (1.2) et rappelons les travaux sur la perception du timbre vocalique (1.3).

Le **chapitre deux** compare les systèmes du tchèque et du français des points de vue :

- grapho-phonique (2.1),
- segmental (2.2), en précisant également les études antérieures concernant les valeurs formantiques des voyelles,
- et suprasegmental (2.3).

Le **chapitre trois** traite de l'apprentissage phonétique et nous y développons :

- les objets généraux d'un apprentissage phonétique (3.1),
- l'apprentissage des voyelles en général (3.2), puis des voyelles du français dans des études antérieures (3.3),
- et enfin les différentes théories de l'apprentissage phonétique (3.4) dont le SLM que nous testons dans cette thèse.

Le **chapitre quatre** traite de l'enseignement de la phonétique. Nous y passons en revue :

- les outils de l'enseignement phonétique (4.1) dont l'API, les triangles vocaliques, les modèles articulatoires de la production des sons de la parole et le spectrogramme en temps réel,
- la notion de la « norme phonétique » (4.2) nécessaire pour quantifier les écarts des apprenants,
- les méthodes d'enseignement des LE (4.3),
- et l'évaluation du niveau phonétique (4.4).

Le **chapitre cinq** introduit la notion d' « accent étranger » en développant les facteurs qui agissent sur le niveau phonétique :

- l'âge au début de l'apprentissage (5.1),
- la qualité et la quantité d'input (5.2),
- l'utilisation relative de la L2 par rapport à la L1 (5.3),
- le nombre d'années de résidence dans le pays où l'on parle la L2 (5.4),
- le genre de l'apprenant (5.5),
- le type d'apprentissage (supports écrits/ oraux) (5.6),
- les facteurs personnels comme l'empathie ou d'autres éventuels (5.7),
- le jugement de la part des auditeurs natifs (certains seraient plus tolérants que d'autres) (5.8)
- et enfin les propriétés phonétiques et phonologiques des sons de la LE par rapport à celles des sons de la LM (5.9).

En 5.10 les conséquences de l'accent étranger sont abordées.

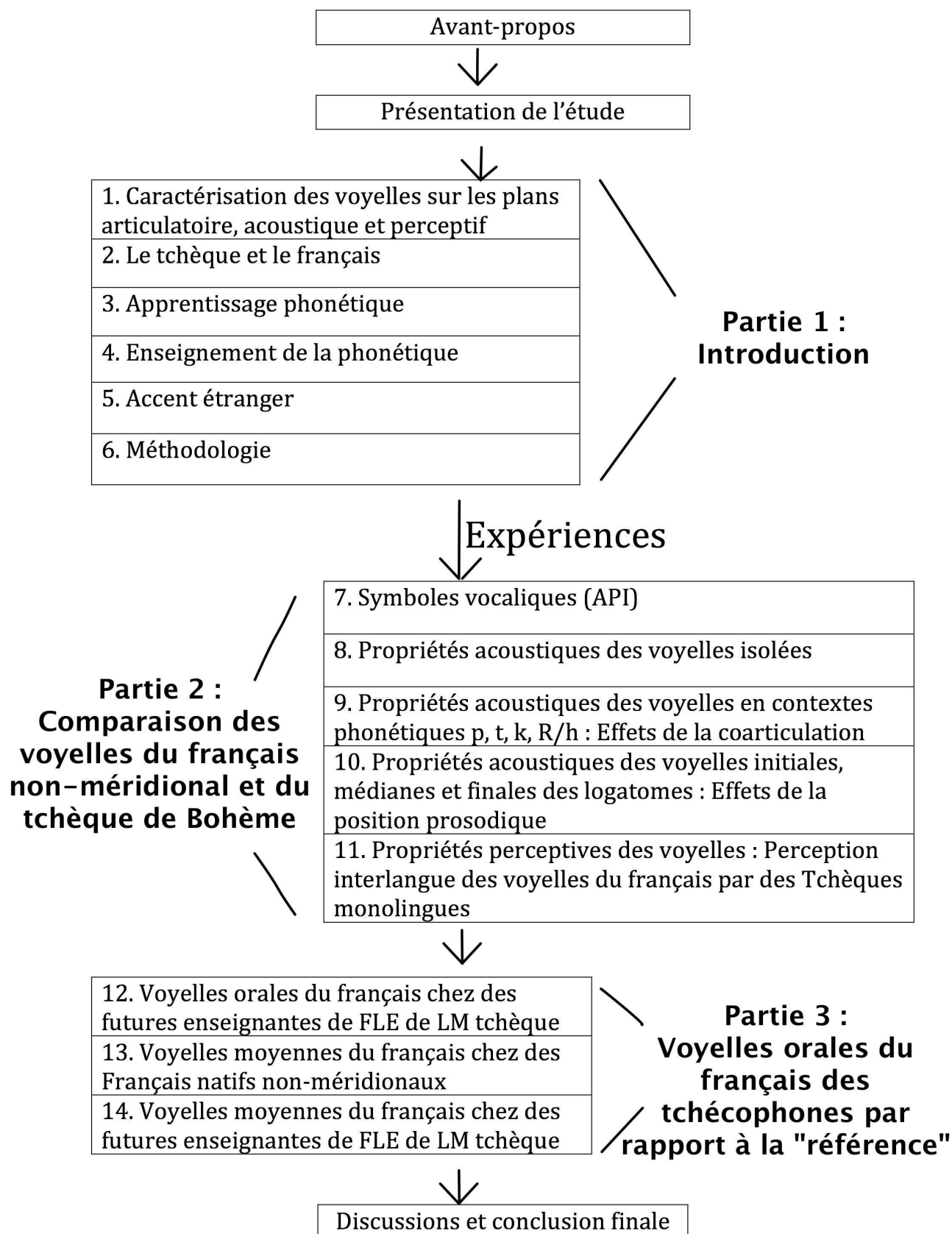


Figure 1 : Structure de la thèse

Le **chapitre six** présente la méthodologie du travail :

- d'abord les locutrices (6.1),
- puis le corpus (6.2),
- les conditions de la passation de l'expérience (6.3),
- la base de données (6.4),
- le traitement des données (6.5),
- le logiciel VisuVo, que nous avons élaboré dans le cadre de la thèse, permettant la visualisation des résultats (6.6),
- l'interprétation des résultats acoustiques (6.7),
- la construction des tests de perception et leur interprétation (6.8),
- et le choix du modèle d'apprentissage phonétique (5.9).

La deuxième partie (chapitres 8 – 11) expose les résultats de la comparaison phonétique des voyelles du français non-méridional et du tchèque de Bohème.

Dans le **chapitre sept**, nous comparons les symboles phonétiques de l'API utilisés pour retranscrire les voyelles :

- en français (7.1),
- en tchèque (7.2).

Dans le **chapitre huit**, nous étudions les propriétés acoustiques des voyelles isolées du français et du tchèque. Nous développons notamment :

- les valeurs formantiques des cibles (8.1),
- la focalisation formantique (8.2),
- les indices acoustiques des contrastes vocaliques (8.3),
- la similarité acoustique des voyelles isolées du français et du tchèque, définie en termes d'écarts types des moyennes formantiques (8.4).

Le **chapitre neuf** s'intéresse aux propriétés acoustiques des voyelles produites en contextes symétriques p, t, k, R. Nous observons :

- les effets de coarticulation à partir d'un modèle articulatoire (9.1),
- les effets de coarticulation en français à partir de logatomes CVCVCVC où C représente les consonnes p, t, k, R (9.2),
- les effets de coarticulation en tchèque à partir de logatomes CVCVCVC où C représente les consonnes p, t, k, h (9.3),
- les différences entre les effets de la coarticulation en français et en tchèque (9.4),
- la similarité acoustique des voyelles du français et du tchèque en contextes symétriques p, t, k, définie en termes d'écarts types des moyennes formantiques (9.5).

Le **chapitre dix** aborde les propriétés acoustiques des voyelles initiales, médianes et finales des logatomes en étudiant les effets de la position prosodique. Nous nous y intéressons à :

- la réduction phonétique (variation de la durée et des valeurs formantiques) en français (10.1),
- la réduction phonétique (variation de la durée et des valeurs formantiques) en tchèque (10.2).
- Une conclusion partielle se trouve en 10.3.

Le **chapitre onze** traite de la perception interlangue des voyelles françaises par des Tchèques monolingues dans le but de rendre compte de la similarité perceptive des voyelles du français aux voyelles du tchèque en :

- isolation (11.1),
- contextes symétriques p, t, k (11.2).

La deuxième partie se termine par une conclusion partielle.

La troisième partie (chapitres 12 – 14) est consacrée à la production et à la perception des voyelles orales du français par de futures enseignantes de FLE de LM tchèque que nous comparons à la « référence » (qui est représentée dans notre thèse par des Français natifs non-méridionaux). Nous y mettons l'accent sur les voyelles moyennes.

Le **chapitre douze** examine la fiabilité des prédictions basées sur le SLM concernant les difficultés que rencontreront les Tchèques apprenant le français en quantifiant leurs écarts phonétiques dans la production des :

- voyelles *nouvelles* (12.1),
- voyelles *similaires* aux voyelles tchèques (12.2),
- voyelles en contexte uvulaire (RVR) qui n'existe pas en tchèque (12.3).

Il liste ensuite les catégories vocaliques qui ne sont pas maîtrisées par les futures enseignantes (12.4).

Nous exposons une conclusion partielle en 12.5.

Le **chapitre treize** examine le statut des voyelles moyennes du français chez des Français natifs non-méridionaux. Il traite de leur :

- perception (en 13.1),
- et de leur production en (13.2).

Une conclusion partielle conclut en (13.3).

Le **chapitre quatorze** s'intéresse enfin à la maîtrise des voyelles moyennes du français chez dix futures enseignantes individuellement :

- en perception (14.1),
- puis en production (14.2).

Une conclusion partielle du chapitre douze se trouve en (14.3.).

La thèse se termine par les discussions et la conclusion finale.

Première partie : Introduction

La première partie (chapitres 1 – 6) introduit le cadre de la thèse. Nous rappelons les théories, les hypothèses, les notions et les études antérieures sur lesquelles nous nous appuyons dans l'interprétation des résultats. Nous décrivons ensuite le protocole expérimental.

1. Caractérisation des voyelles sur les plans articulatoire, acoustique et perceptif

Le jeune professeur trouve son premier outil dans les dimensions articulatoires des segments ... [mais ce n'est pas] un luxe de demander au professeur d'avoir également quelques notions acoustiques sur les sons du langage ... [et enfin] le professeur devra connaître la théorie de la production du langage.

(Delattre, 1960, p. 483 - 485)

Résumé : Dans le **premier chapitre**, nous rappelons les bases de la théorie acoustique de la parole qui explique la formation des sons en termes d'une ou de plusieurs sources et de filtres supra-glottiques (Fant, 1960). Elle élucide ainsi la relation entre la forme du conduit vocal et les formants des sons résultants.

Nous décrivons ensuite la coarticulation des sons, qui est inhérente à la parole et résulte d'un chevauchement des gestes articulatoires lors de la production de sons qui se succèdent (Fant, 1973). L'ampleur de la coarticulation dépend de plusieurs facteurs comme par exemple le débit et le style de parole, le nombre de voyelles dans le système phonologique et la position prosodique. Selon Lindblom (1963), les voyelles brèves (produites en débit rapide) subissent davantage d'influence du contexte consonantique environnant et donc une plus grande réduction phonétique que les voyelles longues (produites en débit lent) s'éloignant ainsi de la cible articulatoire sous-jacente. Selon Manuel and Krakow (1984), les voyelles des langues avec un petit nombre de qualités vocaliques (comme le tchèque) ont des plages de variation acoustique plus grandes que les voyelles des langues avec un grand nombre de timbres (comme le français). Enfin Delattre (1969) démontre que les voyelles inaccentuées dans les langues à accent mobile (comme l'anglais) sont généralement acoustiquement plus réduites que les voyelles inaccentuées dans les langues à accent fixe (comme le français).

Nous nous intéressons ensuite à la perception du timbre vocalique. Alors que les modèles statiques décrivent les qualités vocaliques par la valeur centrale de leurs formants et sont particulièrement adaptés aux langues ayant des voyelles monophthongues de forte tension, comme le français, le modèle dynamique prend en compte la fréquence et la durée non seulement du noyau vocalique mais également des parties environnantes. Ce modèle est adapté aux langues possédant des voyelles diphtongues ou diphtonguées comme l'anglais. Un modèle statique est suffisant pour le français.

1.1 Théorie acoustique de la production de la parole

En 1942, Chiba and Kajiyama ont publié le livre intitulé *The Vowel: Its Nature and Structure*, précurseur de la théorie moderne sur les rapports entre production des sons et leurs caractéristiques acoustiques. La *théorie source-filtre* (« source-filter theory ») a été largement développée dans les travaux de Fant (1960) et de Stevens (1998). Ces auteurs introduisent la physique dans le domaine de la phonétique et ils font le lien entre l'anatomie du conduit vocal (articulatoire), les fréquences de

résonance (l'acoustique) et la perception des voyelles (psychoacoustique et psycholinguistique), (Arai, 2004).

1.1.1 Théorie source-filtre

Selon la *théorie source-filtre*, unanimement acceptée, les sons de la parole sont formés par filtrage d'une ou de plusieurs sources, comme illustré à la Figure 2.

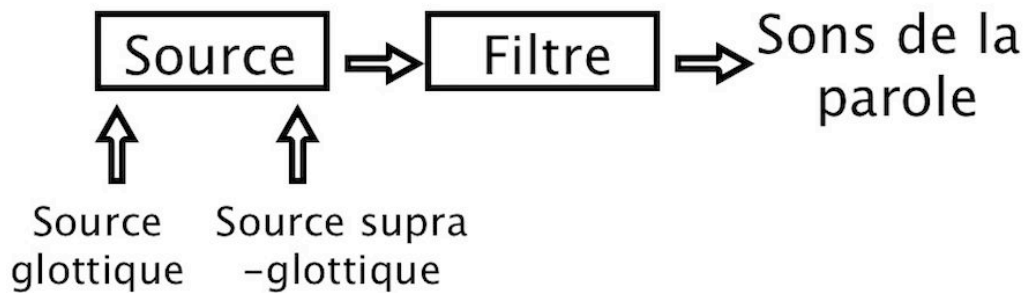


Figure 2 : Formation des sons de la parole selon la théorie source-filtre.

1) Source glottique pour les sons voisés (dont les voyelles)

L'air est expulsé des poumons lorsque la pression trans-glottique est positive, c'est à dire lorsque la pression sous-glottique est supérieure à la pression intra-orale. Le flux d'air rencontre la glotte, l'espace entre les deux plis vocaux. Les plis vocaux, s'ils sont mollement resserrés, font obstacle au passage de l'air et se mettent à vibrer et créent une source glottique. Le flux d'air continu est alors modulé dans le cas des sons voisés. La source glottique voisée (dit aussi bourdonnement) est composée d'un fondamental (qui est inversement proportionnel à la durée de chaque cycle d'ouverture et de fermeture des plis vocaux) et d'harmoniques qui sont des multiples du fondamental.

2) Filtre supraglottique

Le timbre particulier de chaque voyelle est généré ensuite au niveau de la cavité supra-glottique par filtrage acoustique du signal glottique. Les caractéristiques résonancielles du filtre dépendent de la forme du conduit vocal. Selon la configuration du conduit vocal, des zones d'harmoniques de la source glottique sont renforcées (ces zones correspondent aux formants) ou au contraire, affaiblies. Un formant correspond à un pic spectral dans le signal de la parole, c'est-à-dire à une zone d'harmoniques dont l'amplitude est renforcée. La forme générale du conduit vocal est essentiellement due à la position de la langue, l'articulateur principal, et aussi à la configuration des lèvres, du voile du palais et de la hauteur du larynx. La langue crée un resserrement à un endroit le long du conduit vocal, plus ou moins antérieur ou postérieur sur l'axe horizontal et plus ou moins étroit sur l'axe vertical (Stevens, 1998; Vaissière, 2006) : le resserrement est plus étroit (ou complet) pour les consonnes que pour les voyelles. Les voyelles antérieures sont produites avec une constriction au niveau du palais dur alors que les voyelles postérieures correspondent à une constriction dans la partie centrale et postérieure du conduit, de la partie vélaire à la partie pharyngale. La voyelle [œ], quant à elle, est produite sans constriction.

Lorsque la constriction n'est pas très étroite, comme pour certaines voyelles, les cavités en avant et en arrière de la constriction communiquent, le couplage acoustique entre les deux parties est

important, et par conséquent elles ne résonnent pas de façon indépendante. Les formants sont alors dus à l'ensemble du conduit vocal et il est généralement impossible de relier un formant à une seule cavité (Martin, 2008).

Lorsque les voyelles sont produites avec une forte constriction, ou si les deux cavités ont des aires très différentes, il est possible d'associer formants et cavité en avant ou en arrière de la constriction (Fant, 1960). Ainsi pour les voyelles extrêmes, comme le [i] cardinal (forte constriction antérieure), le [ɑ] postérieur cardinal (forte constriction au niveau du pharynx), ou le [u] cardinal (avec deux fortes strictions, au niveau du voile du palais et des lèvres), ou encore les autres voyelles cardinales [y, o, ɔ], il est possible de lier au moins un des formants à une cavité donnée. A ce propos, Stevens (1998, p. 142 - 143) s'exprime ainsi :

There are, however, many situations of interest in which the coupling is small, and approximations to the natural frequencies of the entire configuration can be obtained by estimating the frequencies of the individual components.

Les valeurs des formants pour un phonème donné varient dans une certaine plage. Certaines variations s'expliquent par des facteurs non-linguistiques, tels que l'anatomie du locuteur, l'état mental et physique du locuteur, son sexe et son âge, et par des facteurs linguistiques liés à l'origine socio-économique et géographique du locuteur (Duchet, 1981; Yang, 1996). L'anatomie du conduit vocal et la distribution entre la cavité palatale et pharyngale diffèrent entre hommes, femmes et enfants (Apostol *et al.*, 2004) : l'abaissement du larynx chez les hommes font que ceux-ci ont une cavité postérieure relativement plus longue que les femmes par rapport à la cavité antérieure ; ainsi les formants essentiellement dus à la cavité postérieure sont proportionnellement plus bas chez les hommes que chez les femmes et les enfants. D'autres variations sont conditionnées par la distribution des phonèmes dans la langue, le contexte phonétique ou le débit et style de parole. Pour les locuteurs de la même communauté linguistique, ces variations conditionnées sont assez régulières (Meunier, 2001) et par conséquent font partie du système phonologique de la langue (Guimbretière, 1994).

1.1.2 Lien articulatoire-acoustique dans la modélisation des voyelles

Si Chiba et Kajiyama, puis Fant et Stevens ont pu poser les fondements de la théorie acoustique de la parole, les progrès sont dus initialement aux modèles mécaniques du conduit vocal, élaborés sur la base de données physiologiques, telles que la radiographie, palatographie et laryngoscopie, puis aux modélisations basées sur l'informatique.

Le nombre de tubes droits strictement nécessaires pour modéliser correctement une voyelle varie. Il est possible de modéliser la forme du conduit vocal pour produire une voyelle par un (la voyelle neutre [œ]), deux (les voyelles [ɑ], [i] ou [y]) ou quatre tubes (la voyelle labiale [u]) (Stevens, 1998). Il existe des modèles plus complexes comme celui de Maeda (1982) qui utilise sept paramètres articulatoires (ce modèle est plus largement décrit dans le chapitre 4 qui traite des outils dans l'enseignement de la phonétique).

La voyelle [œ] peut être modélisée par un seul tube, fermé à la glotte et ouvert aux lèvres (Figure 3, page 28) long de 17,5 cm pour un homme. Les formants correspondent alors à des résonances quart d'onde de ce tube. La formule qui permet de calculer une résonance quart d'onde est la suivante : $F = c/4L * (2n+1)$, où F correspond à la résonance, c représente la vitesse de l'air et L la longueur du conduit vocal. L'espacement des formants F1-F2-F3 pour la voyelle [œ] en français est équidistant (Kamiyama and Vaissière, 2009). Un allongement de ce tube va faire baisser tous les formants et on obtient le timbre [ø].

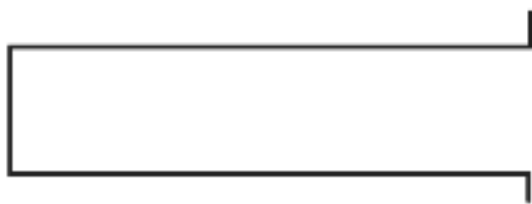


Figure 3 : Modèle avec un tube pour [œ], avec la glotte à gauche et les lèvres à droite. Adaptation selon (Fant 1960)

La voyelle pharyngale [ɑ] peut être modélisée par deux tubes (Figure 4) : un tube pharyngal très étroit, fermé-ouvert qui part de la glotte, comprend la constriction et correspond à une résonance quart d'onde qui détermine le premier formant (ou le second si le tube postérieur est plus court que le tube antérieur). Le tube buccal fermé-ouvert est celui entre la constriction et les lèvres, il est large et présente une résonance quart d'onde, souvent à l'origine du deuxième formant. Pour résonner à la même fréquence, les deux cavités doivent être de longueur effective similaire.

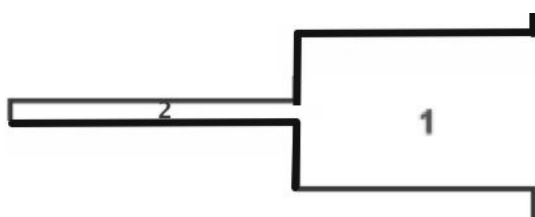


Figure 4 : Modèle de deux tubes pour [ɑ] avec la glotte à gauche et les lèvres à droite. 1 = cavité buccale, 2 = cavité pharyngale. Adaptation selon Fant (1960)

Dans les cas des voyelles antérieures très fermées [i, y], la valeur très basse de F1 est une résonance de Helmholtz. Cette résonance résulte de la combinaison d'une grande cavité (ici la cavité postérieure) avec un goulot étroit. La fréquence peut être calculée avec la formule $FH = \frac{c}{2\pi} \sqrt{\frac{A}{L \cdot V}}$. La fréquence F1 diminue quand le volume de la cavité postérieure V (cm³) augmente, quand la longueur L (cm) du goulot augmente et quand le volume A (cm²) du goulot diminue. Les deux tubes, antérieur et postérieur, (fermé-fermé et ouvert-ouvert) engendrent des résonances mi-onde ($F = \frac{nc}{2L}$). F2 dépend du tube le plus long et F3 du tube le plus court : le F2 de [i] est une résonance demi-onde de la cavité postérieure qui est plus longue que la cavité antérieure, calculée avec la formule $F = \frac{nc}{2L}$, et F3 une résonance demi-onde de la cavité antérieure (la plus courte des cavités). Dans le passage de [i] à [y], il se produit un changement d'affiliation. La cavité antérieure se prolonge par l'avancement des lèvres, devient plus longue que la cavité postérieure, et détermine désormais le F2. En revanche, la cavité postérieure se raccourcit et détermine le F3. Les voyelles [i, y], modélisées avec deux tubes, sont démontrées à la Figure 5. Afin d'obtenir un F3 maximale élevé dans la cas de [i], il faut réduire au maximum la longueur de la cavité antérieure, et donc faire une constriction la plus avancée et la plus étroite possible dans la partie avant du conduit vocal et d'étirer les lèvres. Afin d'obtenir un F3 maximale bas dans le cas de [y], il faut allonger au maximum la longueur effective de la cavité antérieure, et donc avancer et arrondir les lèvres et éventuellement reculer la constriction et l'élargir.

Pour modéliser la voyelle [u], les quatre tubes du modèle de Fant sont nécessaires car deux constriction se produisent, au niveau vélaire et au niveau labial. Les quatre tubes correspondent respectivement à la cavité pharyngale, à la constriction vélaire réalisée par le dos de la langue, à la cavité buccale et à la constriction labiale.

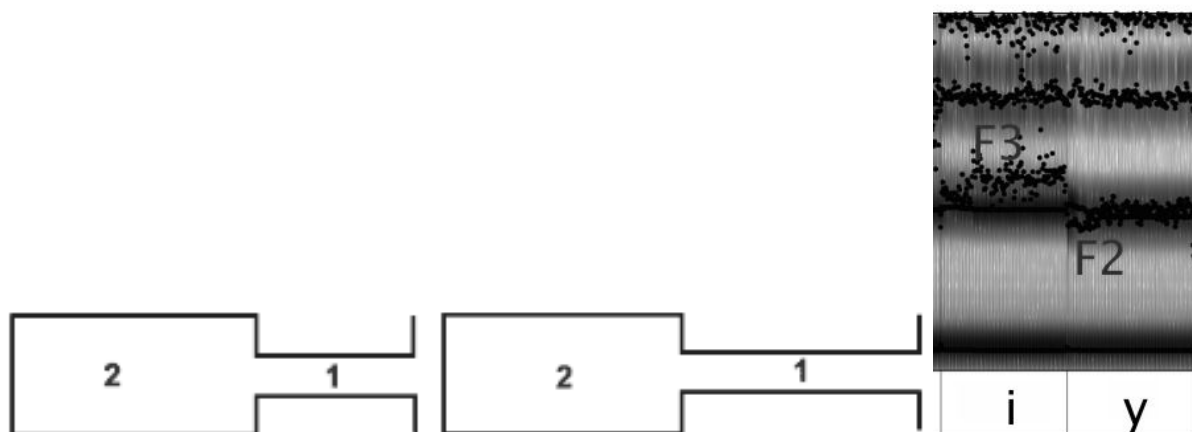


Figure 5 : Modèle de deux tubes pour [i] (à gauche) et [y] (au milieu), avec la glotte à gauche et les lèvres à droite. 1 = cavité buccale, 2 = cavité pharyngale. Adaptation selon (Fant 1960). Changement d'affiliation de cavités que l'on peut voir sur le spectrogramme du [i] et [y] (à droite)

Les deux premiers formants sont dus à une résonance de type Helmholtz, le troisième formant serait dû à la résonance demi-onde du tube de constriction centrale et le quatrième formant dépend de la résonance demi-onde du tube buccal (Maeda and Carré, 1996). L'énergie pour [u] se concentre dans la région (F1F2), et les troisième et quatrième formants, peu intenses, ont un poids perceptif négligeable.

Dans le modèle à quatre tubes de Fant (1960), les paramètres utilisés pour faire varier la longueur et l'aire de chaque tube sont au nombre de trois et ils correspondent à la position de la constriction entre la glotte et les lèvres, au degré de la constriction et/ou à l'arrondissement/ étirement des lèvres. Les nomogrammes de Fant (voir Figure 6, page 30) illustrent les conséquences acoustiques d'un changement de l'un des ces trois paramètres articulatoires en termes d'évolution des cinq premiers formants. Les fréquences des formants dus à une cavité diminuent lorsque la longueur de cette cavité augmente. La longueur de la cavité antérieure diminue quand la cavité postérieure s'allonge et donc les formants dus à chacune de ces cavités évoluent en sens inverse, si la configuration des lèvres reste identique. Dans le cas des formants liés à la cavité antérieure, leur valeur diminue au fur et à mesure que la constriction s'éloigne des lèvres, c'est-à-dire lorsque la langue recule et/ou les lèvres s'allongent (par arrondissement ou protrusion). Nous avons indiqué par des flèches à la Figure 6 les endroits de convergence acoustique où les résonances des cavités en avant et en arrière de la constriction se croisent (appelés également les points focaux). Les convergences de F1/F2 se trouvent dans la région près de la glotte alors que les convergences de F2/F3 ou F3/F4 se produisent dans la région en avant du conduit vocal.

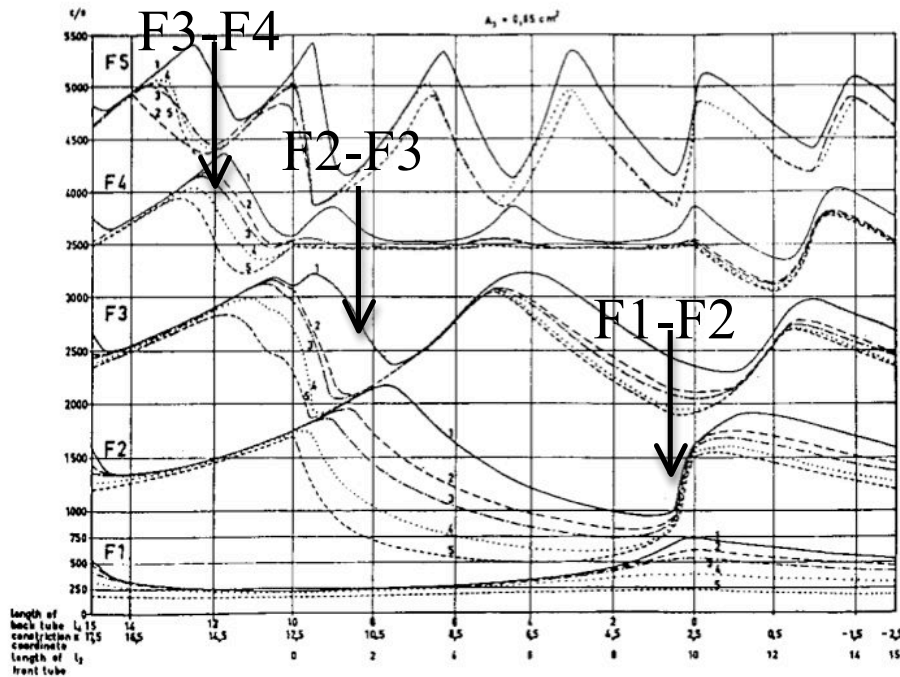


Figure 6 : Modèle à quatre tubes et nomogrammes de Fant (1960) représentant la glotte à droite, les lèvres à gauche. Les flèches indiquent les points focaux

1.1.3 Théorie Quantique

Les points focaux ont fait émergé un certain nombre de théories de production et perception de la parole, dont la *Théorie Quantique* et la *Théorie de la Perturbation*. Selon ces théories développées par Stevens (1989), les endroits de convergence acoustique correspondent à des zones stables, appelées « quantiques », où un mouvement articulaire de faible amplitude n'a pas de grande incidence acoustique, contrairement aux zones instables où un léger déplacement d'un organe articulaire peut entraîner un changement important au niveau acoustique. Les zones stables sont propices à la production de certaines voyelles car elles ne nécessitent pas de gestes articulatoires précis. Ces convergences donnent naissance à des voyelles dites « focales » (Badin *et al.*, 1990) qui se caractérisent ainsi par un rapprochement de formants. Notons que Stevens a essentiellement monté sa théorie à partir des deux premiers formants des voyelles de l'anglais alors que Vaissière (2011) a remarqué que F3 était très sensible au moindre mouvement lorsque les deux premiers formants étaient stables (la littérature américaine prête de façon générale une attention très réduite au troisième formant). Vaissière insiste sur l'importance des formants supérieurs pour la définition des voyelles focales antérieures du français [i, y]. Les notations que l'auteure propose pour les voyelles cardinales [i, y, u] et la pharyngale [ɑ] en isolation, sont les suivantes :

[i] = prépalatal (\uparrow F3F4) 3200 Hz

[y] = (F2F3) 1900 Hz

[u] = (\downarrow F1 \downarrow F2) 400 Hz

[ɑ] = \uparrow (F1F2) 1000 Hz

La valeur (en Hz) indique la fréquence à laquelle convergent les deux formants (pour un locuteur masculin). \uparrow indique que le formant atteint la valeur la plus élevée possible alors que \downarrow

indique la valeur la plus basse possible. Le formant souligné est celui qui est principalement dû à la cavité antérieure.

Les voyelles [o, ɔ] également considérées comme focales en français, gardent les formants (F1F2) proches par le recul de la langue (diminution de la longueur de la cavité postérieure) compensé par un degré d'arrondissement aux lèvres moindre (diminution de la longueur de la cavité antérieure). Vaissière (2011) fait correspondre à la Figure 7 les points de convergence du nomogramme (a) aux rapprochements formantiques effectifs sur des spectrogrammes des voyelles focales [i, y, u, o, ɔ, a] produites par un Français natif (b).

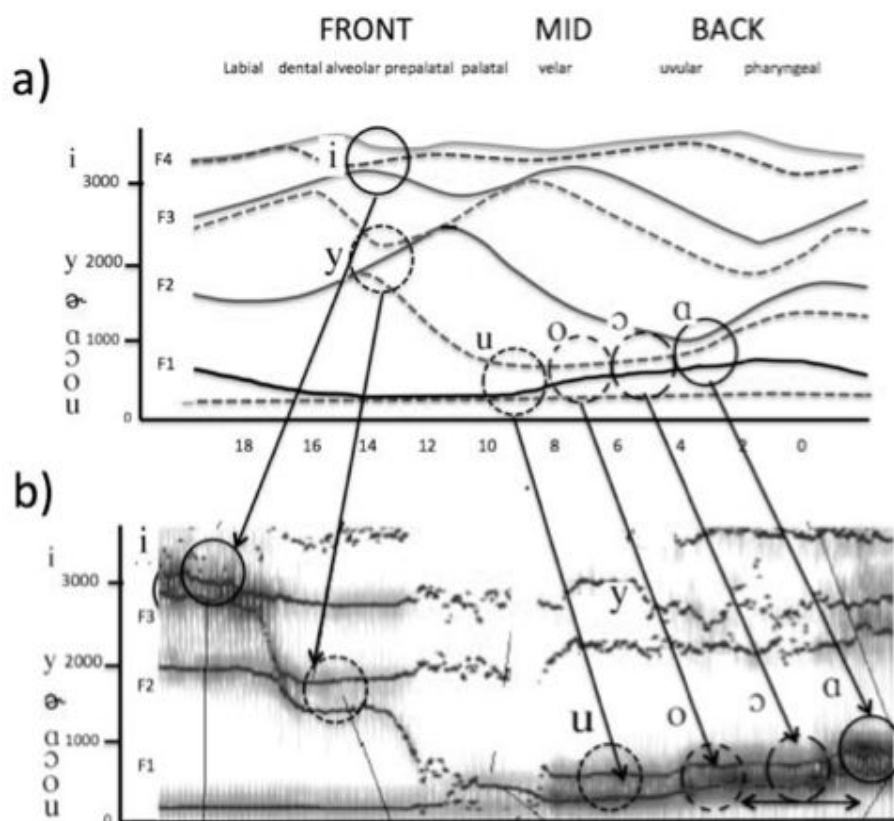


Figure 7 : Les points de convergences formantiques dans un nomogramme (image du haut) et spectrogramme correspondant des 6 voyelles focales du français [u, o, ɔ, a, i, y], prononcées par un locuteur Français natif, selon (Vaissière, 2011)

Enfin, les régions instables seraient celles des consonnes, où un petit mouvement articulatoire déclenche un effet acoustique important (sous la forme de transitions).

1.2 Coarticulation

Les voyelles et les consonnes n'apparaissent que très rarement en isolation. La plupart du temps, elles s'influencent mutuellement (Wioland, 2005). Au cours de la production de sons successifs, les gestes articulatoires se chevauchent du fait de l'inertie variable des organes et il n'y a pas de frontière exacte entre deux sons successifs (Fant, 1973). Ce phénomène est décrit en termes de coarticulation qui est donc inhérente à la parole et « *c'est pourquoi, la prise en compte de variations dues à la coarticulation est désormais incontournable dans la description de la parole.* » (Meunier, 2007). Öhman (1966) constate que l'effet du contexte phonétique en anglais est peu visible dans la partie stable de la voyelle, alors qu'il est évident à son début et à sa fin dans les transitions formantiques. Durand (1985), lui, mesure le degré de coarticulation dans la partie stable des voyelles du français.

Un certain nombre de paramètres peuvent influencer le degré de coarticulation observée dans la partie centrale d'une voyelle :

1. **Coordination des organes de la parole** : Le degré de coarticulation est déterminé du point de vue physiologique par le chemin que doivent parcourir les articulateurs pour passer d'un son à l'autre (Stevens and House, 1963). Si les sons sont articulés dans la même partie du conduit vocal ou avec le même degré d'arrondissement des lèvres, la langue et les lèvres ont une courte distance à parcourir et la déformation des cibles due à la coarticulation est moindre. Si en revanche les articulateurs doivent beaucoup bouger pour la réalisation des sons successifs, le changement de configuration des cavités de résonance provoque des variations notables de la fréquence des formants (Howie, 2001).

2. **Nombre et distribution des voyelles présentes dans le système vocalique** : La coordination des gestes articulatoires est le résultat d'un long apprentissage et la façon de coarticuler est fonction de la phonologie de chaque langue : le degré de coarticulation entre consonne et voyelle devrait être logiquement plus important dans les langues avec un inventaire vocalique réduit, car plus une langue a de qualités vocaliques dans son système, plus la précision articulatoire est élevée afin de maximiser le contraste entre les différents timbres (Manuel, 1990; Manuel and Krakow, 1984). Cette tendance n'est néanmoins pas observée dans toutes les langues. Ainsi Maddieson and Wright (1996) observe qu'en amis, langue parlée en Taïwan, les valeurs formantiques F1 et F2 des seules trois voyelles [i, a, u] du système vocalique présentent une petite dispersion autour de la moyenne, malgré leur insertion dans des contextes consonantiques divers.

L'effet du contexte consonantique est difficilement séparable dans la parole naturelle des autres types de réductions occasionnées notamment par :

1. **Débit et style de parole** : En débit lent ou dans la production des voyelles longues, les valeurs de formants prises à la moitié de la durée de la voyelle (qui est la partie stable) se rapprochent des caractéristiques acoustiques de la cible vocalique (Stevens, 1998). En revanche, en cas de débit rapide, on a souvent affaire à la non réalisation des valeurs prototypiques de la cible vocalique (que Lindblom (1963) appelle le « target undershoot »). Le style de parole détermine également le degré de coarticulation : en parole hypo-articulée on s'éloigne de la cible acoustique alors qu'en parole hyper-articulée, on s'en rapproche (Lindblom, 1990).
2. **Accent lexical** : Dans les langues à fort accent lexical, il existe une tendance à neutraliser les oppositions vocaliques dans les syllabes atones par rapport aux syllabes accentuées où les contrastes se réalisent. Cette tendance semble être plus forte dans les langues à accent mobile, comme l'anglais où la réduction vocalique est très importante (Delattre, 1969; Fokes and Bond, 1989) que dans les langues à accent fixe, comme le français et le tchèque (Delattre, 1969; Volin and Studenovsky, 2007).

Outre le degré de coarticulation qui varie d'une langue à l'autre, le sens de la coarticulation, rémanente ou anticipatoire, peut également dépendre de la langue. La coarticulation est dite anticipatoire si un segment de la chaîne influence le segment qui le précède et la coarticulation est dite rémanente si un segment de la chaîne influence le segment qui le suit.

En français, il existe un lien fort entre l'attaque et le noyau et la coarticulation est essentiellement anticipatoire (Gottfried, 1984). Par exemple, les consonnes qui sont suivies de voyelles labiales présentent déjà un degré d'arrondissement important (Delattre, 1953; Flege, 1988b). Hirsch *et al.* (2004) ont étudié dans les suites sV la possibilité de percevoir certaines propriétés acoustiques des voyelles arrondies du français [y, u, o, ø] durant la consonne fricative précédente [s]. Le résultat a montré qu'il était possible de catégoriser les voyelles [y, u] durant la fricative [s] par le bruit de friction qui descendait au niveau du troisième formant de la voyelle arrondie suivante. En revanche, la

perception anticipatoire n'était pas observée dans le cas de [s] suivi de [o, ø] car le bruit de friction était stable et ne visait pas la valeur basse du troisième formant de la voyelle labiale suivante.

Dans les variations acoustiques dues au contexte phonétique, **tous les formants ne sont pas affectés avec la même amplitude.**

1. **Premier formant** : Le formant F1 connaît des transitions de moins grande envergure que les autres formants, du moins en anglais (Hillenbrand *et al.*, 2001; Stevens, 1998; Stevens and House, 1963). En français, dans un corpus non contrôlé, la valeur du premier formant des voyelles ouvertes généralement baisse lorsque celles-ci sont encadrés de consonnes (Meunier *et al.*, 2006), mais pas dans le contexte de la consonne uvulaire (Vaissière, 2009). Il existe ensuite une tendance à augmenter le F1 pour les voyelles fermées (Delattre, 1969, Wioland, 2005).
2. **Deuxième formant** : En français et en anglais, la coarticulation est particulièrement visible au niveau du deuxième formant, quand les consonnes et les voyelles en séquence ne partagent pas le même lieu, antérieur ou postérieur, de constriction (Manuel, 1990), ou le même degré d'arrondissement des lèvres.
 - Dans le cas des consonnes labiales, la constriction se fait aux lèvres et les résonances correspondant à la cavité en arrière de la constriction sont basses (Stevens, 1998). La cavité antérieure est absente (Johnson, 1997). Durant la production de la consonne, la langue est libre à prendre la position nécessaire pour l'articulation de la voyelle subséquente. Si la consonne labiale est adjacente à une voyelle postérieure arrondie (avec les deux premiers formants bas), les changements formantiques sont minimes. Si, en revanche, la consonne labiale est au contact d'une voyelle antérieure étirée, un abaissement particulier du ou des formants associés à la cavité antérieure est observé. Selon Stevens (1998, p. 342), « ...when a labial consonant precedes a front vowel... the second (or sometimes the third) natural frequency for the vowel configuration is associated primarily with the front cavity... ».
 - Dans le cas des consonnes dentales ou alvéolaires, la constriction se trouve à environ 1,5-2,5 cm des lèvres et le corps de la langue est très avancé (la lame ou l'apex s'élève vers le palais dur). Ceci forme une cavité antérieure courte avec une résonance élevée (Stevens, 1998). Les lèvres sont libres et peuvent s'étirer ou s'arrondir selon la voyelle adjacente (Howie, 2001), et si elles s'arrondissent, la cavité antérieure s'allonge et les fréquences des formants dus à cette cavité diminuent. Si la consonne coronale est au contact d'une voyelle antérieure, les mouvements formantiques sont moindres. Si en revanche la consonne coronale est adjacente à une voyelle postérieure, la valeur du deuxième formant de cette dernière naturellement basse augmente en se dirigeant vers le « locus dental » qui se trouverait aux alentours de 1800 Hz en français (Delattre, 1963).
 - Dans le cas des consonnes vélaires ou dorso-palatales, le corps de la langue s'élève contre le palais dur créant ainsi une constriction postérieure (Johnson, 1997). La cavité antérieure s'allonge car la constriction (relativement longue) se trouve désormais à environ 5-6 cm des lèvres et les résonances dues à cette cavité baissent. Si les deux cavités résonnent à la même fréquence, cela provoque une convergence acoustique du deuxième et troisième formant (Stevens, 1998).

Ainsi, la *labialisation* provoque une chute de tous les formants, et dans une plus grande mesure du ou des formants dus à la cavité antérieure. La *palatalisation* fait baisser le premier formant et élève le deuxième formant. Enfin, la *vélarisation* fait baisser d'une manière importante le deuxième formant et augmente le premier formant (Brosnahan and Malmberg, 1970; Vaissière, 2009).

Deux études classiques étudiant les effets acoustiques du lieu d'articulation de la consonne sur le premier et le deuxième formant des voyelles de l'anglais doivent être mentionnées. Stevens and House (1963) ont étudié les formants de huit voyelles orales de l'anglais américain [i, ɪ, ε, æ, u, ʊ, ʌ, ɑ]

produites par trois locuteurs masculins, en contexte nul V (voyelles produites isolément), en contexte hVd et dans des syllabes symétriques həCVC où C correspond à l'une des 14 consonnes de l'anglais qui sont divisées en trois groupes : labiales, post-dentales et vélaires. Ainsi, des valeurs formantiques moyennes pour chacune des huit voyelles sont calculées en fonction du contexte (consonantique ou nul). Le résultat montre que les effets acoustiques systématiques du contexte consonantique sur le premier formant de la voyelle adjacente sont minimes alors qu'ils sont plus ou moins importants dans le cas du deuxième formant qui atteint des valeurs plus centralisées que celles des voyelles prononcées isolément. De cette façon, le F2 des voyelles postérieures dans tout contexte consonantique confondu augmente et celui des voyelles antérieures baisse. Notons que l'anglais ne possède pas de consonnes uvulaires ou pharyngales, qui ont pour effet d'augmenter l'aperture des voyelles environnantes et F1. La Figure 8 montre que les voyelles antérieures qui se trouvent au contact d'une consonne vélaire portent des changements formantiques moindres que quand elles sont au contact des consonnes labiale ou post-dentale. Quand les voyelles postérieures sont coarticulées avec une consonne labiale, le changement du deuxième formant est moindre que quand elles sont en contact d'une consonne post-dentale.

Une deuxième étude, de Hillenbrand *et al.* (2001), corrobore les résultats de Stevens and House (1963). Les auteurs examinent l'effet des contextes consonantiques labial, alvéolaire et vélaire sur les mêmes huit voyelles orales de l'anglais américain que Stevens et House, à savoir le [i, ɪ, ε, æ, u, ʊ, ʌ, ɑ] prononcés par six hommes dans des syllabes symétriques et asymétriques CVC avec sept consonnes initiales et six consonnes finales différentes. Les résultats des contextes symétriques montrent que l'effet acoustique de la consonne sur le premier formant de la voyelle adjacente est minime, mis à part les [ε] et [æ] en contextes alvéolaire et vélaire où le premier formant est plus bas par rapport à la cible. En revanche, cet effet est de nouveau plus ou moins important dans le cas du deuxième formant qui est marqué par des valeurs centralisées. Le deuxième formant des voyelles postérieures, notamment celui de [u], augmente fortement au contact des consonnes alvéolaires et dans une moindre mesure au contact des consonnes vélaires et celui des voyelles antérieures baisse au contact des consonnes labiales (Figure 9, l'image en haut, page 36). Les auteurs examinent également l'effet acoustique de la consonne initiale, qui provoque des changements formantiques plus importants que la consonne finale. Ainsi, les transitions formantiques initiales sont plus importantes (Figure 9, l'image du bas à gauche) que les transitions formantiques finales (Figure 9, l'image du bas à droite). Ce phénomène est particulièrement saillant quand les consonnes de la syllabe CVC sont alvéolaires et la voyelle est [u] ou [ʊ] où la valeur du F2 du début et milieu vocalique est particulièrement élevée.

Il existe d'autres études qui ont comparé les stratégies de coarticulation de langues différentes. Elles ont mis en évidence que les différences sont observables en production (par des mesures de formants dans le cas des voyelles) mais elles se projettent également dans la perception.

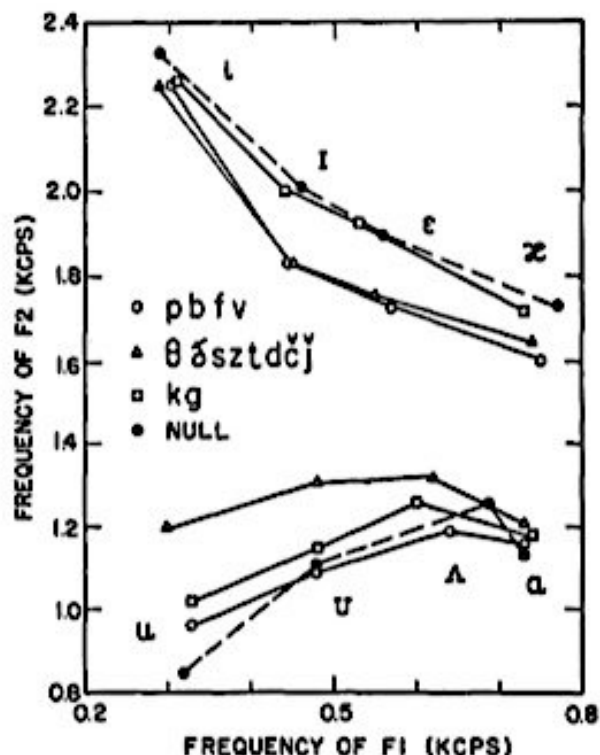


FIGURE 6. Values of F_1 and F_2 in cps for eight vowels plotted to demonstrate the effect of place of articulation of the consonantal context. The contexts are divided into three groups—velars (squares), postdentals (triangles), and labials (circles). Values for vowels in the null environment (solid points) are included for comparison.

Figure 8 : Effet des contextes consonantiques labial, post-dental et vélaire sur les formants F_1 et F_2 de huit voyelles de l'anglais. Selon Stevens and House (1963)

Ainsi Gottfried (1984) compare les formants vocaliques des voyelles de l'anglais de Stevens et House aux formants des voyelles du français [i, e, ε, a, ɑ, u, o, ɔ, y, ø, œ] produites par 4 locuteurs français en isolation et en contexte coronal dans des syllabes tV, tVt et Vt (les valeurs moyennes de F_1 et F_2 se trouvent à la Figure 10, page 37). L'auteur montre que les voyelles du français sont moins affectées par le contexte coronal que les voyelles de l'anglais, notamment le [u] et le [o], dont le F_2 augmente considérablement par rapport à la cible. En français, les voyelles [a] et [ɑ] de même que [ɔ] et [ɔ] sont acoustiquement très proches, en particulier en contexte Vt et en isolation alors que les voyelles [ø] et [œ] sont relativement éloignées.

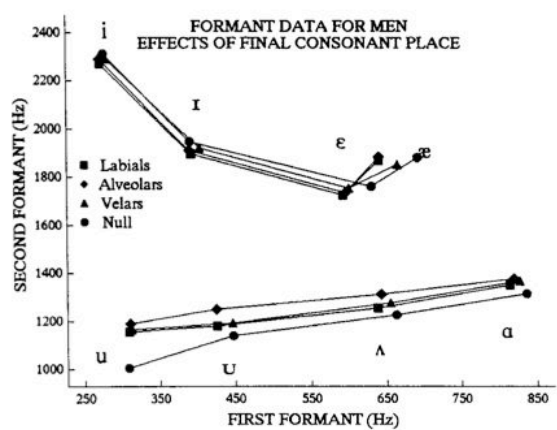
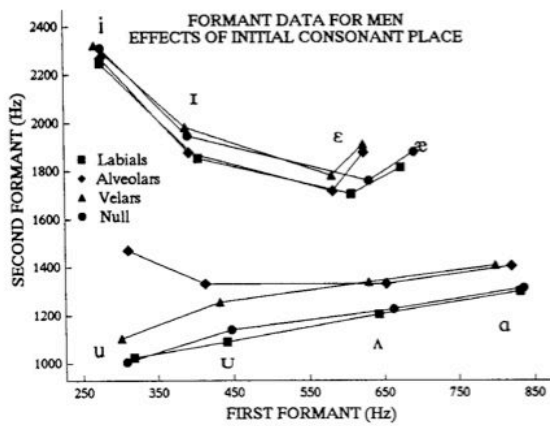
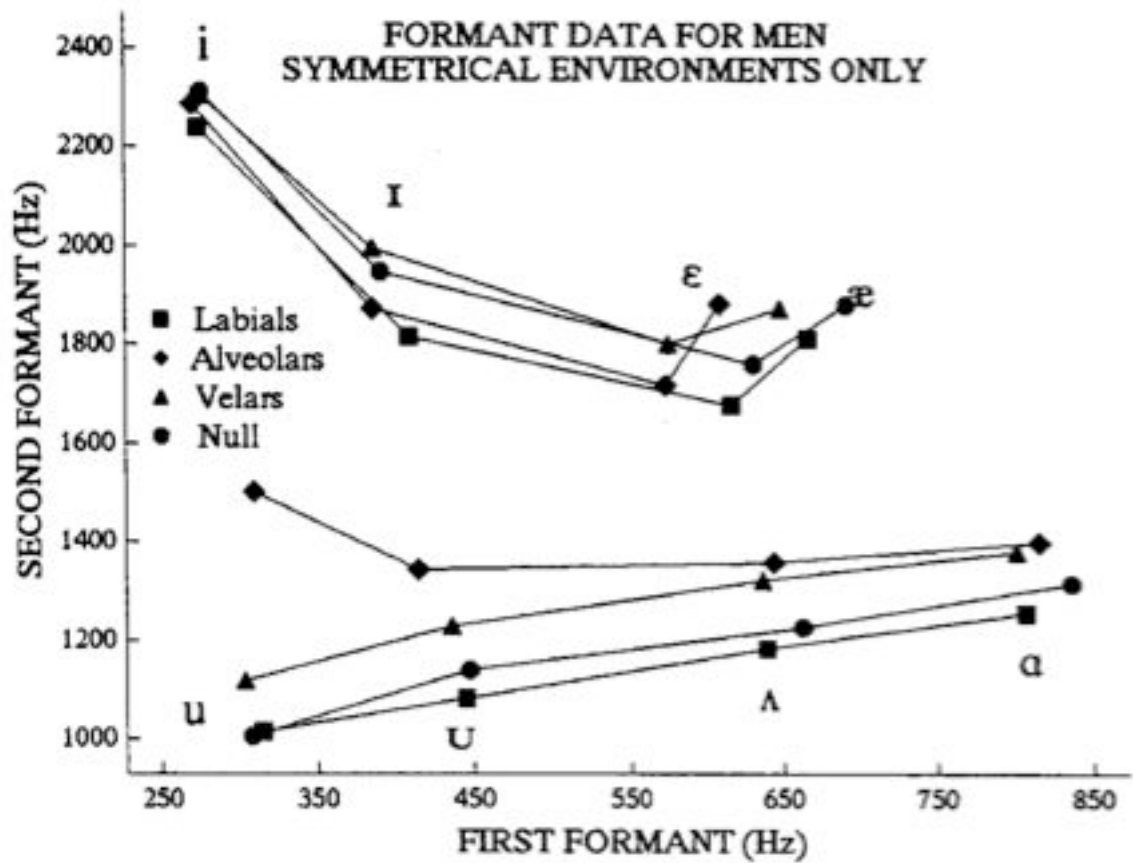


Figure 9 : Influence du contexte consonantique sur les formants F1/F2 (en Hz) des voyelles anglaises prononcées par 6 hommes dont les valeurs sont relevées à la moitié (image du haut) et au début ou à la fin (image du bas, respectivement à gauche et à droite) de la durée vocalique. Selon Hillenbrand *et al.* (2001)

1. Caractérisation des voyelles sur les plans articulatoire, acoustique et perceptif

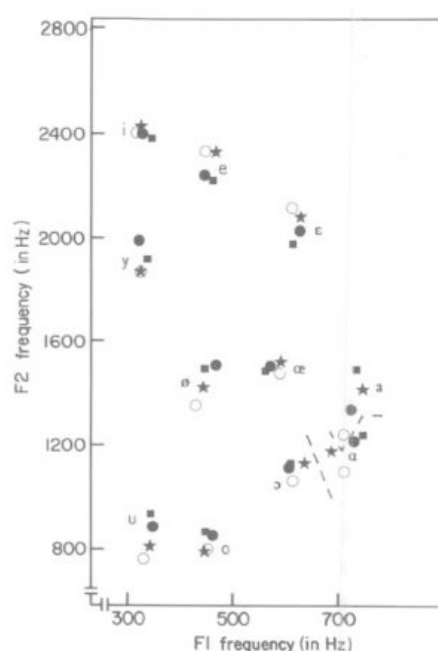


Figure 10 : F1 et F2 moyens (en Hz) des voyelles du français [i, e, ε, a, α, u, o, ɔ, y, ø, œ] (4 locuteurs français natifs*11 voyelles* 2 répétitions) en isolation (rond blanc), en syllabe tVt (carré), en syllabe tV (rond noir) et en syllabe Vt (étoile), selon Gottfried (1984)

Strange *et al.* (2005) ont étudié la similarité acoustique et perceptive de onze voyelles de l'anglais américains [i, ɪ, e, ε, æ, u, ʊ, o, ɔ, ɑ, ʌ] (4 locuteurs natifs* 6 contextes* 3 répétitions) avec 14 voyelles de l'allemand du nord [i, ɪ, e, ε, y, ʏ, ø, œ, u, ʊ, o, ɔ, ɑ, a] (4 locuteurs* 5 contextes* 2 répétitions) ayant été produites dans des syllabes CVC où C correspondent aux occlusives labiales, alvéolaires et vélaires. L'analyse acoustique en termes des formants F1 à F3 a montré qu'en contextes alvéolaire et vélaire, les voyelles antérieures arrondies de l'allemand sont proches des voyelles postérieures de l'anglais, acoustiquement antériorisées (Tableau 1).

NG front vowel	AE back vowels	b-p % of tokens	d-t % of tokens	g-k % of tokens
y	u	0%	100%	75%
ø	u, ou	87	100	100
ʏ	u, ʊ	75	100	100
œ	u, ʌ, u	13	100	100
All 4 NG vowels	AE Back vowels	48%	100%	94%

Tableau 1 : Similarité spectrale des voyelles antérieures arrondies de l'allemand du nord (NG) et des voyelles postérieures de l'anglais américain (AE) en fonction du contexte, selon Strange *et al.* (2005). % indique le pourcentage de voyelles antérieures de l'allemand qui correspondent aux voyelles postérieures de l'anglais dans un contexte donné

De même, les voyelles de l'allemand [ɪ, e, ε, ɔ], généralement considérées comme *identiques* ou "hautement similaires" aux [ɪ, e, ε, ɔ] anglais, sont situées dans l'espace acoustique « plus haut » que leurs homologues allemands. Enfin, la similarité acoustique entre les voyelles des deux langues varie en fonction du contexte consonantique. Dans la tâche d'assimilation perceptive, 48 auditeurs anglais monolingues ont identifié les 14 voyelles de l'allemand au travers de onze voyelles de l'anglais et ils ont noté les stimuli sur une échelle de 1 (« very foreign-sounding ») à 7 (very « English-sounding »). Le résultat montre que les voyelles antérieures arrondies et les voyelles postérieures de l'allemand sont assimilées aux voyelles postérieures de l'anglais avec la même note moyenne. L'assimilation perceptive ne varie pas en fonction du contexte, c'est-à-dire qu'une voyelle de l'allemand est systématiquement assimilée à une seule catégorie anglaise indépendamment du contexte

(Tableau 2). Du plus, l'assimilation perceptive ne peut pas être toujours prédite à partir des données acoustiques. Les auteurs ont alors montré l'importance de comparer les sons des deux langues au niveau acoustique mais également perceptif.

Overall modal responses				Context-specific modal responses					
NG	All five contexts			b-p context		d-t context		g-k context	
	AE	%	(Mdn)	%	(Mdn)	%	(Mdn)	%	(Mdn)
i	i	99	(7)	99	(7)	99	(7)	99	(7)
ɑ	ɑ	75	(5)	81	(5)	79	(5)	65	(5)
u	u	88	(5)	90	(5)	92	(5)	82	(5)
y	u	93	(5)	93	(4)	95	(5)	89	(5)
Y	u	74	(4)	68	(4)	72	(4)	73	(4)
ø	u	43	(5)	47	(4)	42	(5)	41	(5)
œ	Λ	74	(4)	77	(4)	75	(4)	71	(4)
ɪ	ɪ	84	(6)	98	(6)	94	(6)	52	(5)
e	e	64	(6)	63	(5)	60	(6)	60	(5)
ɛ	ɛ	92	(6)	98	(6)	97	(6)	70	(6)
ʊ	ʊ	45	(4)	38	(3)	36	(3.5)	31	(4)
o	ou	69	(5)	65	(5)	72	(5)	68	(5)
ɔ	ɔ	53	(4)	36	(4)	48	(4)	72	(4)
a	ɑ	55	(5)	69	(5)	64	(5)	46	(4)

Tableau 2 : Les motifs d'assimilation perceptive des voyelles de l'anglais américains (AE) en termes de catégories de l'allemand du nord (NG), tout contexte confondu (colonne 3-4), en contextes b-p (colonne 5-6), d-t (colonne 7-8) et g-k (colonne 9-10), selon Strange *et al.* (2005)

Strange *et al.* (2001) se sont intéressés à l'assimilation perceptive des 11 voyelles de l'anglais [i:, ɪ, eɪ, ɛ, æ:, u:, ʊ, ɔ:, ɑ:, ʌ] en termes de 18 catégories du japonais. Ainsi cinq voyelles d'un more /a, e, i, o, u/ de même que de 2 mores /aa, ee, ii, oo, uu/, plus [ei], [ou] et des voyelles palatalisées d'un ou de deux mores [ja(a), ju(u), jo(o)] ont été identifiées par 24 auditeurs japonais monolingues qui ont utilisé une échelle de justesse allant de 1 à 7. Les voyelles de l'anglais avaient été produites par 4 locuteurs américains natifs, dans des syllabes CVC où C correspond aux occlusives labiales, alvéolaires et vélaires. Le résultat montre qu'aucune des onze voyelles de l'anglais n'est systématiquement assimilée à une catégorie japonaise, tout contexte confondu. Ainsi, la plus grande similarité perceptive est établie pour [eɪ] anglais, relié dans 77 % à [ei] japonais alors que la similarité la plus faible est établie pour la voyelle *nouvelle* [æ:], reliée au [aa] japonais dans 32 % des réponses. La catégorisation des voyelles longues (notamment de [i:] et [u:] anglais) en termes de voyelles du japonais dépend fortement du caractère voisé de la consonne en coda. Quand les réponses par les catégories du japonais regroupent les mêmes qualités vocaliques sans prendre en compte la palatalisation et les différentes durées vocaliques, la plus forte similarité perceptive est observée entre [i:] anglais et [i, ii] japonais (assimilation dans 98 % des réponses). A l'inverse, la similarité perceptive la plus faible existe pour les voyelles anglaises considérées par les auteurs comme *nouvelles*. Il s'agit du [ɪ] et du [ʌ], assimilés dans 51 % des réponses respectivement aux [i, ii] et [a, aa, ja, jaa]. Les auteurs ont conclu que la difficulté à identifier et discriminer les voyelles de l'anglais par des japonophones dépend fortement du contexte consonantique dans lequel les voyelles figurent.

Levy (2009) s'est intéressée à l'assimilation perceptive de sept voyelles françaises [y, œ, u, o, i, ɛ, a] en termes de onze catégories anglaises /i, ɪ, e, ɛ, æ, ɑ, ɔ, ʌ, o, ʊ, u/ et du [ju, ɜ] par 13 anglophones américains monolingues. Les voyelles du français avaient été produites par trois Françaises natives en contextes labial rabVp ou alvéolaire radVt et les auditeurs ont utilisé une échelle de justesse de 1 (« foreign sounding ») à 9 (« native sounding »). Le résultat met en évidence que l'assimilation des voyelles antérieures arrondies du français aux voyelles postérieures de l'anglais dépend du contexte (voir Tableau 4).

1. Caractérisation des voyelles sur les plans articulatoire, acoustique et perceptif

AE vowel	Japanese response categories				
	High-front i, ii	Mid-front e, ee, ei	Low a, aa /a, /aa	Mid-Back o, oo, ou, /o, /oo	High-back u, uu, /u, /uu
ii	98	2			
i	81	44	1		4
ej	5	98			
e	2	83	44	<1	<1
æi	1	26	73		
ai		1	81	16	2
A	<1	1	81	39	8
ot			13	82	5
ou	<1		<1	79	20
u			3	15	82
ui				<1	99

Tableau 3 : Les motifs d'assimilation des voyelles de l'anglais américain (AE) aux voyelles japonaises sans prendre en compte leur durée et la palatalisation, tout contexte consonantique confondu, selon Strange *et al.* (2001)

Ainsi [y] est majoritairement assimilé à [ju], dans 80 % des réponses en contexte bilabial et dans 65 % des réponses en contexte coronal. Le [œ] est ensuite assimilé principalement au [u] en contexte bilabial (dans 38 % des réponses) et au [u] en contexte coronal (dans 59 % des réponses). La plus grande similarité est établie entre le [i] français et le [i] anglais (dans 98 % des réponses en contexte bilabial et dans 97 % des réponses en contexte coronal).

No Exp							
PF	AE	%	Med				
i	i	98	8	i	i	97	7
	e	79	7	e	e	76	6
æ	æ	12	6		e	11	6
					æ	10	6
a	ɑ	58	6	a	æ	57	7
	æ	36	6		ɑ	29	7
o	u	62	7		e	12	6
	o	24	7	o	u	43	6
u	u	90	7		o	41	6
	ju	6	4	u	u	84	6
y	ju	80	6		ju	9	4
	i	12	4	y	ju	65	4
œ	u	38	6		u	31	6
	u	34	6	œ	u	59	6
	æ	1	7		o	17	6
					u	17	5

Tableau 4 : L'assimilation perceptive des voyelles du français parisien (PF) en termes de catégories anglaises (AE) en contextes labial (à gauche) et coronal (à droite), selon Levy (2009)

La comparaison des formants vocaliques, sur les seuls plans F1 et F2 (Figure 11, page 43) des voyelles du français (en gras) et de l'anglais a montré que l'assimilation perceptive des voyelles antérieures arrondies du français aux voyelles postérieures de l'anglais ne peut pas s'expliquer entièrement par les données acoustiques. En effet en contexte bilabial, le [y] français est acoustiquement plus proche du [i] que du [u] anglais et [œ] est à mi-chemin entre les voyelles antérieures et postérieures de l'anglais.

Expérimentateur(s)	Type d'étude	Corpus/ Stimuli-réponses		Locuteurs		Auditeurs	Résultat
		1.	2.	1.	2.		
Gottfried (1984)	Acoustique : mesure des formants F1, F2	[i, ɪ, ε, æ, u, ʊ, ʌ, ɑ] anglais en isolation, hVd ou en həCVC où C = labial/ post- dental/vélaire	[i, e, ε, æ, a, u, o, ɔ, y, ø, œ] en français en isolation et en tV, tVt, Vt	3 Américains natifs	4 Français natifs		- voyelles du français moins affectées par le contexte coronal que les voyelles de l'anglais, notamment le [u] et le [o], dont le F2 augmente par rapport à la cible d'avantage en anglais qu'en français
Strange et al. (2005)	Acoustique : mesure des formants F1, F2, F3	[i, ɪ, e, ε, y, ʏ, ø, œ, u, ʊ, o, ɔ, ɑ, ʌ] allemands en bVp, bVt, dVt, gVk, gVt	[i, ɪ, e, ε, æ, u, ʊ, o, ɔ, ɑ, ʌ] anglais en bVb, bVp, dVd, dVt, gVg, gVk	4 Allemands du Nord	4 Américains natifs		- en contextes alvéolaire et vélaire, les voyelles antérieures arrondies de l'allemand sont proches des voyelles postérieures de l'anglais, acoustiquement antériorisées - [ɪ, e, ε, ɔ] allemands réalisés avec F1 plus bas que [ɪ, e, ε, ɔ] anglais - la similarité acoustique entre les voyelles des deux langues varie en fonction du contexte consonantique
	Perception interlangue	Stimuli : [i, ɪ, e, ε, y, ʏ, ø, œ, u, ʊ, o, ɔ, ɑ, ʌ] allemands en bVp, bVt, dVt, gVk, gVt	Réponses : /i, ɪ, e, ε, æ, u, ʊ, o, ɔ, ɑ, ʌ/ anglais (dans des mots hVd)	48 Américains			- voyelles antérieures arrondies et voyelles postérieures de l'allemand assimilées aux voyelles postérieures de l'anglais avec la même note moyenne => l'assimilation perceptive ne varie pas en fonction du contexte (association toujours à la même catégorie anglaise) => l'assimilation perceptive ne peut pas s'expliquer entièrement par les données acoustiques
Strange et al. (2001)	Perception interlangue	Stimuli : [i:, ɪ, eɪ, ε: æ:, u:, ʊ:, o:, ɔ:, ɑ:, ʌ] anglais en CVC où C = labial/ alvéolaire/vélaire	Réponse : /a, e, ɪ, o, u, aa, ee, ii, oo, uu/ et [ei, ou, ʌa(a), ʌu(u), io(o)] japonais	4 Américains natifs		24 Japonais monolingues	- aucune des 11 voyelles de l'anglais systématiquement assimilée à une catégorie japonaise, tout contexte confondu : catégorisation des voyelles longues (notamment [i:] et [u:]) de l'anglais en termes de voyelles du japonais dépend du caractère voisé de la consonne en coda - la plus forte similarité observée entre [i:] anglais et [i, ii] japonais (98 %) - la plus faible assimilation pour les voyelles anglaises considérées comme <i>nouvelles</i> [ɪ] et [ʌ], reliées dans 51 % aux [i, ii] et [a, aa, ʌa, ʌaa] respectivement => l'assimilation perceptive dépend du contexte

1. Caractérisation des voyelles sur les plans articulatoire, acoustique et perceptif

Levy (2009)	Perception interlangue	Stimuli : [y, œ, u, o, i, ε, a] français en contextes labial rabVp ou alvéolaire radVt	Réponse : /i, i, e, ε, æ, a, ɔ, ʌ, o, u, u/ et [u, ɜ-] anglais (dans des mots hVd)	3 Françaises natives		13 Américains monolingues	<ul style="list-style-type: none"> - l'assimilation des voyelles antérieures arrondies du français aux voyelles postérieures de l'anglais dépend du contexte : [y] est assimilé à [u], dans 80 % en contexte bilabial et 65 % en contexte coronal, [œ] assimilé à [o] en contexte bilabial (38 %) et à [u] en contexte coronal (59 %) - la plus forte similarité de [i] français avec [i] anglais, dans 98 % en contexte bilabial et dans 97 % en contexte coronal
	Acoustique : mesure des formants F1, F2	[y, œ, u, o, i, ε, a] français en contextes labial rabVp ou alvéolaire radVt	[i, i, e, ε, æ, a, ɔ, ʌ, o, u, u, ɜ-] anglais en contexte bilabial	3 Américaines natives			<ul style="list-style-type: none"> - en contexte bilabial, le [y] français est acoustiquement plus proche du [i] que du [u] anglais - [œ] est à mi-chemin entre les voyelles antérieures et postérieures de l'anglais <p>⇒ l'assimilation perceptive ne peut pas s'expliquer entièrement par les données acoustiques</p>
Trofimovich <i>et al.</i> (2001)	Perception interlangue	Stimuli : [i, i, u, o, a, ʌ, æ, ε] anglais en hVd, bVd, bVt, lVC et nVC	Réponse : /i, o, i, y, œ, ε, e, ʌ, a, u/ coréennes	3 Américains natifs		10 Coréens monolingues	<ul style="list-style-type: none"> - [i, i, u, o, a, ʌ] assimilées à une catégorie coréenne prédominante - [æ, ε] assimilées à 2 catégories coréennes distinctes /e/ et /ε/ où aucune n'est prédominante - pour [i, i, u, o, a, ʌ] (mais pas pour [æ] et [ε]) l'assimilation varie selon le contexte

Tableau 5 : Etudes antérieures sur les différences de coarticulation entre les langues

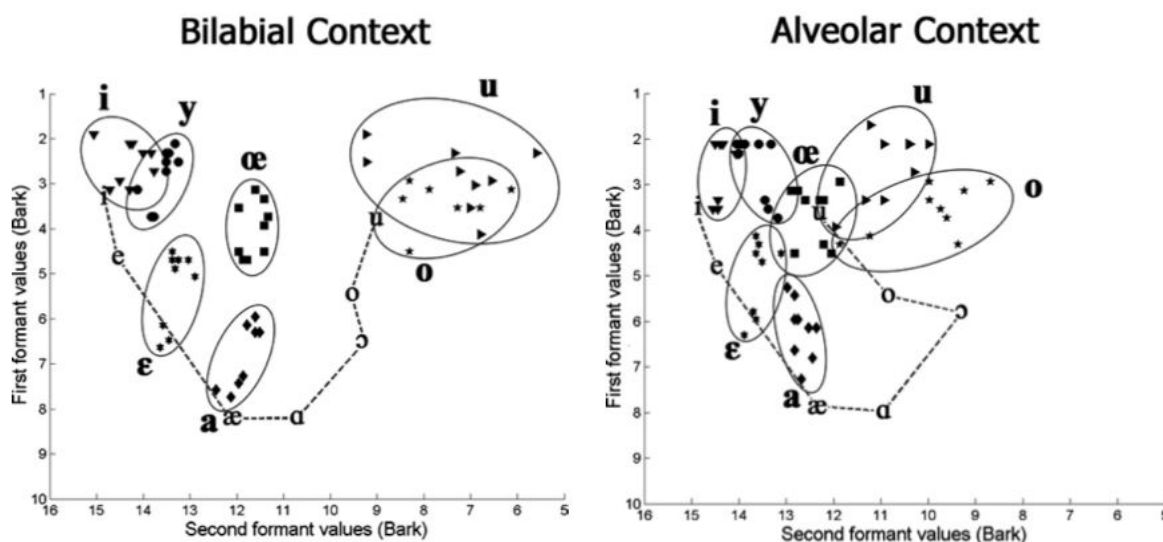


Figure 11 : Triangle vocalique F1/F2 (en Bark) des voyelles du français (en gras) et de l'anglais, produites en contextes bilabiale (à gauche) et alvéolaire (à droite) par 3 locuteurs natifs selon Levy (2009)

Enfin, Trofimovich *et al.* (2001) s'intéressent à la perception interlangue par 10 Coréens monolingues de huit voyelles de l'anglais [i, ɪ, u, ʊ, ɑ, ʌ, æ, ɛ] produites par trois Américains natifs dans cinq contextes différents : en contexte « neutre » hVd et en contextes bVd, bVt, IVC et nVC. Les voyelles des 120 mots ainsi produits ont été assimilées à l'une des dix catégories vocaliques coréennes /i, o, ɪ, y, æ, ε, e, ʌ, a, u/ et évaluées sur une échelle de 1 (faible similarité) à 7 (forte similarité avec la voyelle coréenne). Le résultat montre que les voyelles anglaises [i, ɪ, u, ʊ, ɑ, ʌ] sont assimilées à une catégorie coréenne prédominante alors que les voyelles [æ, ɛ] sont perceptivement reliées à deux catégories coréennes distinctes /e/ et /ɛ/ où aucune n'est prédominante (voir le Tableau 6). Le résultat montre également que dans le cas des six voyelles anglaises [i, ɪ, u, ʊ, ɑ, ʌ], l'assimilation varie selon le contexte consonantique alors que pour le [æ] et le [ɛ], l'assimilation n'est pas influencée par le contexte consonantique. Ainsi, les [i, ɪ, ɑ, ʌ] anglais sont moins fréquemment assimilés aux respectivement /i, i, a, ʌ/ coréens en contexte IVC que dans les autres contextes. Les [i, ʌ, ʊ] anglais sont moins fréquemment assimilés aux /i, ʌ, u/ coréens en contexte nVC que dans les autres contextes. Enfin, les [u, ʊ] anglais sont moins fréquemment assimilés au /u/ coréen en contexte bVd que dans les autres contextes.

Vowel		Modal responses
English	Korean	
/i/	→ /i/	95 (6.0)
	→ /y/	5 (5.3)
/ɪ/	→ /i/	80 (5.6)
	→ /ɪ/	4 (4.9)
/e/	→ /e/	38 (5.2)
	→ /ɛ/	54 (5.3)
/æ/	→ /e/	45 (5.7)
	→ /ɛ/	37 (5.5)
/u/	→ /u/	85 (5.4)
	→ /ɪ/	13 (4.0)
/ʊ/	→ /u/	71 (5.4)
	→ /ɪ/	16 (4.5)
/ɑ/	→ /a/	78 (5.8)
	→ /ʌ/	17 (5.0)
/ʌ/	→ /a/	73 (5.5)
	→ /ʌ/	21 (5.8)

Tableau 6 : Assimilation perceptive des voyelles de l'anglais en termes de voyelles du coréen, selon Trofimovich *et al.* (2001)

Le Tableau 5 récapitule les différentes études qui comparent les stratégies de coarticulation entre deux langues, du point de vue de la production, de la perception ou des deux à la fois.

1.3 Perception du timbre vocalique

Il existe trois modèles de description de la perception du timbre vocalique (Strange, 1989). D'abord *le modèle de la cible vocalique statique*, couramment utilisée dans les manuels d'introduction en phonétique. Ensuite, *le modèle élaboré de la cible vocalique* qui se base également sur des données acoustiques statiques (mesure de la valeur centrale de la voyelle), transformées ensuite en représentations auditives pour mieux refléter leur perception non-linéaire. Enfin, *le modèle de spécification dynamique* utilise l'information dynamique pour décrire les voyelles en contexte qui seraient alors identifiées avec une plus grande précision que les voyelles isolées.

1.3.1 Modèle de la cible vocalique statique

Selon le modèle de la cible statique, la valeur centrale des deux à quatre premiers formants est dite capitale pour la perception du timbre et par conséquent pour la catégorisation de la voyelle (Baken and Orlikoff, 2000).

Dans le cas de rapprochement de deux formants amplifiés (F1-F2, F2-F3, F3-F4) , l'oreille perçoit une seule zone de fréquences pour ces deux formants, perceptivement saillante, plutôt que des formants séparés. Selon Delattre *et al.* (1952, p. 203) :

The ear effectively averages two vowel formants which are close together, receiving from these two formants an impression which is highly similar to that which would be heard from one formant placed at a position somewhere intermediate between them.

Les travaux en laboratoire de Haskins (Delattre *et al.*, 1951; Delattre *et al.*, 1952) initient les recherches sur le rôle des traits acoustiques dans la perception des sons de la parole. Ces travaux sont à la base de la notion de F2' (F2 prime), notion introduite plus tard par Carlson *et al.* (1970), appelé parfois le formant « effectif ». En synthétisant 16 voyelles cardinales à l'aide de l'instrument Pattern Playback, qui convertit les spectrogrammes dessinés à la main en sons, Delattre *et al.* (1952) ont montré la possibilité d'approximer le timbre des voyelles par deux ou une seule résonance principale (dans le cas des voyelles postérieures). Afin de fournir une définition acoustique des voyelles cardinales, ils ont créé 235 patrons formantiques divisés en deux groupes : le premier groupe de patrons fixait une valeur constante de F1 et la valeur de F2 variait par pas de 120 Hz, alors que le deuxième groupe contenait des patrons avec une valeur fixe de F2 et des valeurs de F1 qui variaient par pas de 30 Hz. Les patrons qui permettaient le mieux identifier chacune des 16 voyelles cardinales ont été retenus. Pour montrer qu'il était possible de représenter certaines voyelles par un seul formant, les auteurs écoutaient des sons où les deux formants avaient été remplacés par une seule fréquence qui variait entre la valeur de F1 et F2. Leur résultat montre que les voyelles postérieures (mi)-fermées [u] et [o], ayant les valeurs de F1 et F2 proches, peuvent être caractérisées acoustiquement par un seul formant, proche du F1. Les voyelles [ɔ, ɒ, ɑ, a] peuvent être décrites un terme d'un seul formant, intermédiaire entre F1 et F2. Enfin, les voyelles antérieures nécessitent une description acoustique par deux formants, mis à part le [i] qui, selon les auteurs, peut être représenté par un seul formant proche du « F2 », qui était en fait le F3 (notons que dans les premiers travaux de Haskins, les auteurs ont confondu F3 avec F2 pour la voyelle [i]). Les études plus récentes ont montré que dans le cas de la voyelle cardinale [i], F2' se situe entre F3-F4 alors que pour la voyelle [y], F2' est placé entre F2 et F3. Cette valeur abstraite démontre l'importance des fréquences de résonance supérieures à 2000 Hz dans la perception des voyelles antérieures (Stevens, 1998; Vaissière, 2006).

1.3.1.1 La différence « juste perceptible » (JND, *Just Noticeable Difference*) des formants vocaliques

Lors de la synthèse des voyelles, une certaine liberté dans le choix des fréquences de formants s'offre aux expérimentateurs car les auditeurs sont perceptivement sensibles aux changements de fréquences à partir d'un certain seuil. Cette question est abordée dans la littérature sous le nom de « frequency difference thresholds (df)¹ », « frequency difference limen (DLF)² » ou « just noticeable difference (JND) in frequency³ » (Flanagan, 1972).

Flanagan (1955) s'est intéressé au DLF dans la discrimination de qualités des voyelles synthétisées. Afin de déterminer le DLF du premier formant, il a synthétisé trois sons vocaliques de base, avec un F1 de 300 Hz, 500 Hz et 700 Hz, un F2 stable de 1500 Hz, F3 de 2500 Hz et enfin F4 de 3550 Hz (voir Figure 12). Il a ensuite synthétisé 14 variations de ces trois sons de base, en modifiant uniquement la valeur de F1 par pas de 10 Hz (jusqu'à +/-70 Hz).

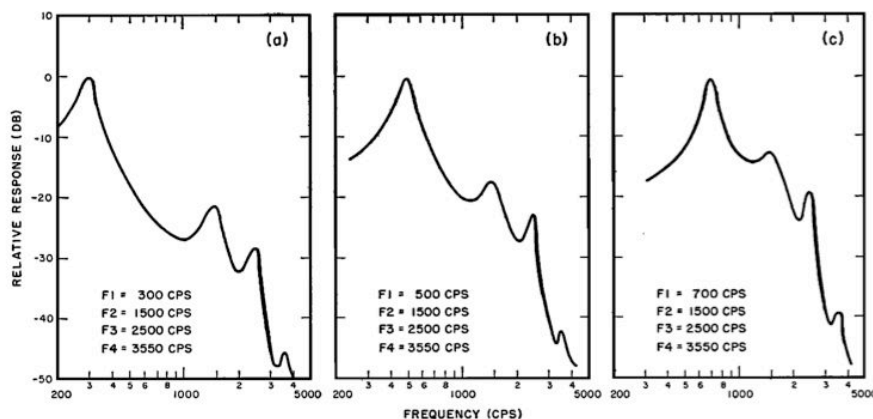


Figure 12 : Les enveloppes spectrales des sons de base utilisés pour tester le « difference limen » de F1, selon Flanagan (1955)

Afin de déterminer le DLF du second formant, Flanagan a synthétisé de nouveau trois sons de base avec un F1 stable de 500 Hz, un F2 de 1000, 1500 et 2000 Hz, un F3 stable de 2500 Hz et F3 de 3550 Hz (voir Figure 13). Il a de nouveau créé 14 variantes de chaque son de base en modifiant cette fois-ci la valeur du deuxième formant par pas de 25 Hz (jusqu'à +/- 175 Hz). Notons que les sons vocaliques de base ne correspondent pas à des voyelles précises ce qui ne pose pas de problème à priori car selon l'auteur (p. 617) : « *Quality DL's can be determined for standard sounds placed anywhere in the F1-F2 plane and not necessarily corresponding to any English vowel phoneme at all.* ».

¹ Nous traduisons : « seuil de la différence perçue »

² Nous traduisons : « seuil de la différence perçue »

³ Nous traduisons : « la différence de fréquence juste perceptible »

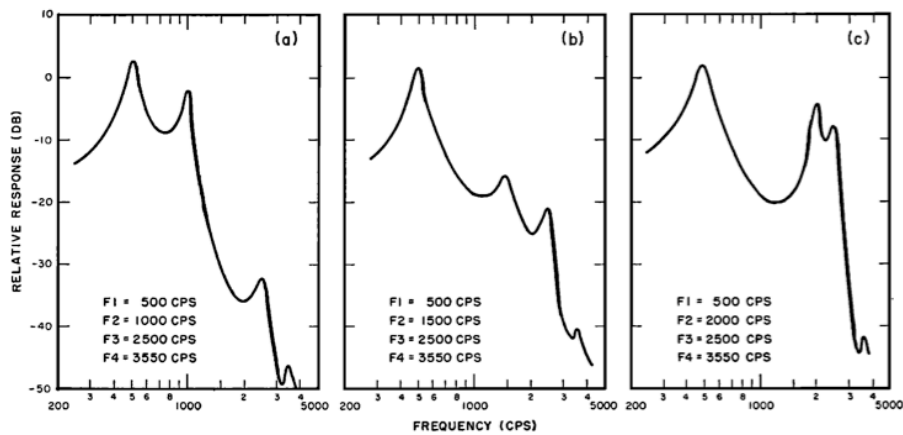


Figure 13 : Les enveloppes spectrales des sons de base utilisés pour tester le « difference limen » de F2, selon Flanagan (1955)

L'auteur a ensuite construit un test de discrimination AB présentant 20 paires de sons dont 14 paires comportaient des sons physiquement différents et six paires contenaient des sons *identiques*. Le temps entre deux stimuli était de 500 ms. Chaque stimulus a été jugé 20 fois et les résultats montrent que le DLF des premier et deuxième formants se trouve entre 3 à 5 % (voir Tableau 7 pour les chiffres exactes).

For the F1 tests:			
	F1 of standard (cps)	-DL (cps)	+DL (cps)
(a)	300	17	12
(b)	500	25	27
(c)	700	27	19
For the F2 tests:			
	F2 of standard (cps)	-DL (cps)	+DL (cps)
(a)	1000	20	50
(b)	1500	45	75
(c)	2000	90	20

Tableau 7 : Le « difference limen » de F1 et F2, selon Flanagan (1955)

Flanagan (p. 616) précise que les DLF établis sont applicables dans la discrimination de différence de qualités vocaliques mais qu'ils peuvent être différents dans le cas d'identification des phonèmes : « *A criterion involving identification of the phoneme would in general lead to much larger formant DL's.* ».

1.3.2 Modèle élaboré de la cible vocalique

Le modèle plus élaboré de la cible vocalique a ensuite montré l'importance perceptive des distances entre les différents formants ainsi qu'entre le premier formant et la fréquence fondamentale, plutôt que l'importance des valeurs formantiques absolues (pour une revue de la littérature, voir Hillenbrand and Gayvert (1993). De cette façon, afin de mieux représenter les distances perçues entre les sons, il est possible d'effectuer d'abord une transformation des fréquences selon l'échelle de Hertz en « bandes critiques » (Bark) qui tient compte de la perception humaine non-linéaire. La formule proposée par Zwicker and Fastl (1990) pour calculer les valeurs Bark est la suivante (où F correspond à la fréquence (en Hertz) :

$$\text{Bark} = 13 \times \arctan(0.76 \times F) + 3.5 \times \arctan\left(\frac{F}{7.5}\right)^2$$

Dans l'étude de Chistovich and Lubilnskaya (1979), l'importance d'un seul pic de fréquence pour la perception des voyelles focales est expliquée par l'effet du centre de gravité qui est calculé à partir de la distance tonotopique entre deux formants rapprochés de grande amplitude. Si deux formants se trouvent proches, ne dépassant pas la distance critique (qui est de 3-3,5 Bark pour F1 et F2), les auditeurs perçoivent un seul formant qui se trouve à mi-chemin entre F1 et F2, et dont la position exacte dépend de l'amplitude relative des deux formants proches. Si la distance critique est dépassée, les auditeurs choisissent alors un formant qui correspond à F1 ou F2 mais sans passer par une fréquence intermédiaire.

Syrdal and Gopal (1986) proposent une normalisation auditive basée sur l'étude de la perception des voyelles de l'anglais, telles qu'elles ont été définies par Peterson and Barney (1952). En calculant les distances tonotopiques entre F1-F0 et F3-F2, les auteurs trouvent une forte corrélation entre la différence de F1-F0, en Bark, et l'aperture, et entre F3-F2, et la perception de l'antériorité/ postériorité. Ainsi les auteurs proposent que les voyelles avec une distance tonotopique F1-F0 inférieure à trois Barks sont perçues comme hautes, alors que lorsque la distance est supérieure à trois Bark, elles sont perçues comme basses. De même, si la distance tonotopique F3-F2 est inférieure à trois Bark, la voyelle est alors perçue comme antérieure alors que si cette distance est supérieure à trois Bark, la voyelle est perçue comme postérieure. Il est à noter que F0 varie beaucoup dans la parole continue, et donc la distance tonotopique F1-F0 aussi. D'autres auteurs (Fant, 1983; Ménard *et al.*, 2010) relient l'antériorité/ postériorité perçue plutôt à la distance tonotopique entre F2-F1.

Notons que toute normalisation permet de ne garder que les caractéristiques acoustiques pertinentes du point de vue linguistique en réduisant la variabilité inter-locuteur (Yang, 1996).

1.3.3 Modèle de spécification dynamique

Enfin, le modèle de spécification dynamique de W. Strange met en évidence l'importance de l'information dynamique, c'est-à-dire des fréquences et de la durée de toutes les parties de la voyelle (onglide = transition formantique initiale, noyau et offglide = transition formantique finale) dans la perception des voyelles en contexte qui seraient en effet des caractéristiques intrinsèques des voyelles et non pas des consonnes environnantes. Strange *et al.* (1983) ont montré qu'il est possible d'identifier les voyelles de l'anglais en contexte dont on a remplacé la partie centrale par un silence alors qu'il est impossible de les identifier lorsque les transitions sont retirées. Strange and Bohn (1998) ont ensuite montré que la durée des trois parties vocaliques joue également un rôle important dans la perception des voyelles de l'allemand, notamment pour distinguer les voyelles moyennes et basses.

Le modèle dynamique, qui prédit une meilleure identification des voyelles en contexte par rapport aux voyelles isolées, est particulièrement adapté à l'anglais se caractérisant par une forte diphtongaison vocalique, et également à l'allemand mais il n'est pas attesté en français. En effet Gottfried (1984) a effectué un test d'identification et de discrimination par 16 Français natifs et huit apprenants anglophones de onze voyelles orales du français [i, e, ε, a, α, u, o, ɔ, y, ø, œ], produites par quatre locuteurs français natifs en isolation et en syllabe tVt, tV et Vt. Le résultat se trouve à la Figure 14 et il montre que les voyelles sont plus faciles à identifier quand elles se trouvent en isolation (20 % d'erreurs chez les natifs) ou qu'elles sont précédées par l'occlusive coronale tV (17,5 % d'erreurs chez les natifs) que lorsqu'elles se trouvent en contexte symétrique tVt ou suivies de l'occlusive coronale Vt.

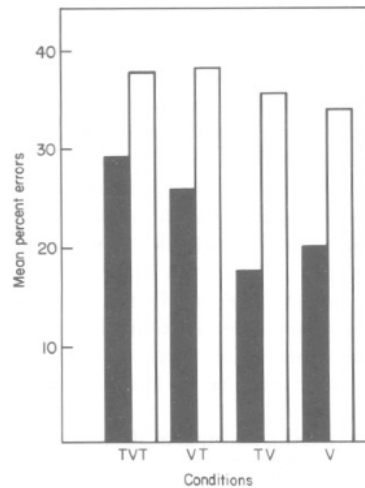


Figure 14 : Le pourcentage moyen d'erreurs d'identification des voyelles du français selon le contexte syllabique auprès des auditeurs français natifs (en noir) et des anglophones américains parlant français (en blanc), selon Gottfried (1984)

Le résultat du test de discrimination catégorielle (voir Figure 15) révèle également une facilité à discriminer les voyelles isolées avec 12,7 % d'erreurs chez les natifs du français par rapport aux voyelles en contextes tVt, discriminées avec 21,9 % d'erreurs par les Français natifs.

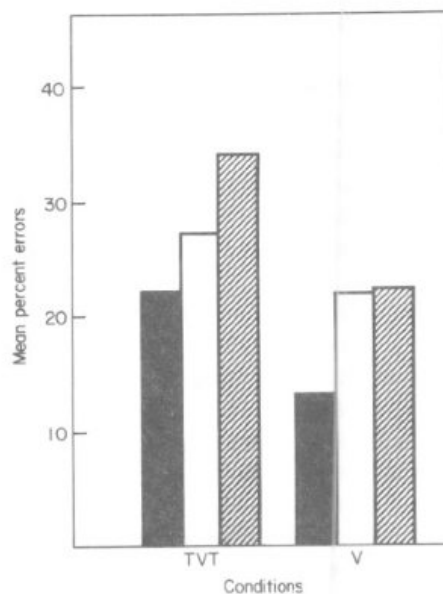


Figure 15 : Le pourcentage moyen d'erreurs de discrimination des voyelles du français selon le contexte syllabique auprès des auditeurs français natifs (en noir) et des anglophones américains parlant français (en blanc) et des anglophones américains monolingues (avec rayures), selon Gottfried (1984)

Ainsi, la perception des voyelles du français est particulièrement bien élucidée par la théorie de la cible vocalique statique alors que la perception des voyelles de l'anglais est mieux définie par le modèle de spécification dynamique.

2. Le tchèque et le français

*Kolik jazyků znáš, tolikrát jsi člověkem.*⁴

Tomáš Garrigue Masaryk

Résumé : Les systèmes segmental et supra-segmental du français et du tchèque sont comparés dans le **chapitre deux**. Le français (langue romane) et le tchèque (langue slave) sont des langues à accent lexical fixe. En français, l'accent réalisé est plutôt post-lexical que lexical et il se réalise sur la syllabe finale d'un mot phonétique alors qu'en tchèque, l'accent lexical est fixé sur la syllabe initiale d'une mesure rythmique (Rigault, 1970). Les deux langues montrent une préférence pour les syllabes ouvertes quoique le français en possède davantage (Dankovicova and Dellwo, 2007). Le système phonologique du français standard actuel comprend dix voyelles orales /i, e, ε, a, u, o, ɔ, y, ø, œ/ et trois nasales /ã, ê, õ/ (ces dernières ne faisant pas l'objet de notre étude). Celui du tchèque correspond traditionnellement à dix voyelles monophthongues (cinq timbres avec une différence de durée) /i, i:, e, e:, o, o:, u, u:, a, a:/ et à trois diphtongues /au, eu, ou/ (Ludvikova and Kraus, 1966). Cependant, le timbre des voyelles hautes du tchèque de Bohême évolue au point que la plupart des auteurs actuels distinguent entre un [i] bref et un [i:] long et certains opposent un [ɔ] bref au [u:] long. De même, la voyelle moyenne antérieure est généralement décrite avec le symbole [ε] et la voyelle moyenne postérieure est parfois notée avec le symbole [ɔ] (Palková, 1997). Dans la partie phonétique de cette thèse, nous désignerons les voyelles par les symboles de l'API qui semblent les plus proches.

La spécificité du système vocalique du français réside dans la série des voyelles antérieures arrondies, des voyelles nasales et des voyelles moyennes à deux degrés d'aperture et celle du tchèque dans la série des voyelles diphtongues et dans l'opposition de voyelles brèves et longues. Le système consonantique du français possède 17 consonnes et 3 semi-consonnes alors que celui du tchèque comporte 25 consonnes. Enfin alors que le tchèque détient une écriture quasi phonétique, le code grapho-phonique du français est opaque.

Le tchèque est une langue indo-européenne du groupe slave occidental. Il est parlé par 10 millions de Tchèques en Bohême et Moravie, les deux régions de la République tchèque et par environ 2 millions d'immigrants qui se trouvent principalement aux Etats-Unis, Canada et dans quelques pays d'Europe (Simackova *et al.*, 2012).

Du point de vue grammatical, le tchèque est une langue inflectionnelle, ce qui signifie que « *les relations grammaticales y sont exprimées par un changement de la structure interne des mots et aussi par des morphèmes autonomes.* » (Ašic, 2008, p. 29). Les morphèmes sont alors soit dérivationnels, servant à dériver de nouveaux mots, soit flexionnels. La flexion nominale du tchèque est riche car elle indique non seulement le genre (masculin, féminin et neutre) et le nombre (singulier et pluriel), comme en français, mais également l'un des sept cas des noms, des adjectifs, des pronoms et des numéraux (Tahal, 2010). Grâce à cette riche morphologie nominale, l'ordre des mots est plus libre en tchèque qu'en français. La flexion verbale indique la personne, le nombre, le temps (passé, présent et futur), le mode, la voix et l'aspect des verbes (perfectif et imperfectif).

Le français est une langue indo-européenne du groupe gallo-roman, résultant de l'évolution du latin, parlée par 140 millions locuteurs (Walter, 1988). Il est notamment parlé en France, en Suisse, en

⁴ Nous traduisons « Un homme vaut autant d'hommes qu'il connaît de langues. »

Belgique, au Canada, dans les anciennes colonies africaines ainsi qu'en Asie et dans les îles de l'Amérique centrale (Dohalska and Schulzova, 2003).

Le français est une langue inflectionnelle, comme le tchèque (Ašic, 2008). Alors que la flexion nominale est moins riche du fait de l'absence de la déclinaison (les cas du latin ont été remplacés par l'utilisation des prépositions en français moderne), la flexion verbale est très développée. Les temps verbaux sont nombreux et correspondent au présent, au passé simple, au passé composé, à l'imparfait, au passé antérieur, au plus-que-parfait, au futur simple et au futur antérieur (Dubois and Lagane, 1995). L'ordre des mots de type SVO (sujet-verbe-objet), fixe depuis la disparition des cas, permet d'exprimer les relations syntaxiques (Flaux, 1993).

2.1 Codes grapho-phoniques

Le tchèque s'écrit au moyen de l'alphabet latin, tout comme le français. Il se sert en plus de certains signes notés au-dessus du graphème pour indiquer sa prononciation. Ainsi le signe *čárka* « petit trait » signale la longueur vocalique et le signe *háček* « petit crochet », ajouté au-dessus des consonnes <t>, <d>, <z>, <s>, <c>, <n>, <r> et la voyelle <e>, indique la « mouillure ». L'écriture du tchèque est phonétique (Tahal, 2010), mis à part en cas de deux consonnes adjacentes, l'une sourde et l'autre sonore où la deuxième impose son trait de voisement à la première (assimilation régressive) et des consonnes sonores qui se dévoient en finale de mot.

L'orthographe du français est plus complexe car la relation entre la graphie et la phonie est irrégulière et opaque (Harmegnies *et al.*, 2005) pour les consonnes et davantage encore pour les voyelles (Léon, 1993). L'écrit nous renseigne sur l'étymologie du mot : les voyelles qui s'écrivent au moyen de digraphes ou trigraphes proviennent des voyelles diphtongues ou triptongues (Lambert-Drache, 1997). Alors que ces sons ont commencé à disparaître à partir du 12^e siècle, aboutissant vers un système des voyelles monoptongues, la graphie n'a pas beaucoup évolué depuis le 18^e siècle (Léon, 1993).

2.2 Systèmes segmentaux

Le phonétisme du français peut être décrit par les trois caractéristiques suivantes : le mode tendu, antérieur et croissant (Delattre, 1953). D'abord, le français exige une grande tension musculaire des organes de la parole et Wioland (2005, p. 73) précise à ce propos que « *En réalité tout relâchement articulaire des articulations CV en syllabe finale est source d'incompréhension.* ».

Le mode croissant est lié avec la tendance ouvrante avec la majorité de syllabes de type CV (Wioland, 2005). La voyelle influence de manière importante l'articulation de la consonne qui précède. Du fait de cette structuration syllabique prédominante et de la forte cohésion entre l'attaque et le noyau, Delattre (1953) propose que les voyelles françaises sont perceptivement plus proéminentes et psychologiquement plus importantes que les consonnes. Dans le cas des syllabes fermées CVC, le coda est affaibli et si possible rattaché, par des phonèmes de liaison et d'enchaînement, à la voyelle de la syllabe suivante.

L'articulation des sons du français est ensuite caractérisée comme majoritairement antérieure : les voyelles palatales, représentant deux tiers du système vocalique, sont composées de deux séries, trois voyelles étirées [i, e, ε] et trois voyelles arrondies [y, ø, œ] (Léon and Léon, 2007). Puisqu'en français l'attaque consonantique prend les traits phonétiques du noyau vocalique, les consonnes sont fortement palatalisées en présence de voyelles antérieures. De même, les consonnes suivies de voyelles arrondies (environ un tiers du système vocalique), prennent la forme de ces dernières ; elles sont ainsi articulées avec la forme arrondie des lèvres. L'antériorisation des sons a débuté au premier siècle, aboutissant à la disparition du timbre [u] du système vocalique du français qui n'a réapparu qu'au 12^e siècle (Vaissière, 1996).

Enfin, les treize voyelles du français contemporain ainsi que le grand nombre de consonnes voisées donnent une caractéristique particulièrement sonore au système français (Léon and Léon, 2007).

En tchèque, contrairement au français, les caractéristiques articulatoires et acoustiques des consonnes sont relativement stables et indépendantes du contexte car elles varient peu en fonction des voyelles subséquentes (Palková, 1997). Ainsi, la palatalisation ne constitue pas un trait du phonétisme tchèque actuel. L'aspiration des consonnes, comme en français, n'a pas lieu. En revanche, l'attaque vocalique est dure, souvent accompagnée d'un coup de glotte. La majorité des syllabes est de type CV.

2.2.1 Systèmes vocaliques

2.2.1.1 Inventaire vocalique du tchèque

Le système phonologique du tchèque correspond traditionnellement à cinq voyelles monophthongues avec une différence de durée /a, a:, e, e:, i, i:, o, o:, u, u:/ (Komarek, 1982; Kucera and George, 1968; Ludvikova and Kraus, 1966) ainsi qu'à trois voyelles diphtongues /au, eu, ou/. Les auteurs actuels soulignent la différence de timbre dans la paire *i/i:* en stipulant que « *[i] bref est légèrement plus ouvert que [i:] long fermé* » (Palková, 1997). Une différence acoustique concerne également les voyelles de la paire *u/u:*. Cette différence pourrait être phonétiquement pertinente car elle dépasse, selon Skarnitzl et Volin (2012), au niveau des deux premiers formants la *différence juste perceptible* (JND) de 5 %. En revanche, les différences des formants dans les paires *ε/ε:*, *o/o:* et *a/a:* sont minimales car inférieures à 5 % (Skarnitzl and Volin, 2012). Par ailleurs les voyelles brèves seraient trois à quatre fois plus fréquentes que les voyelles longues.

Les voyelles monophthongues rencontrent des différences dans leur réalisation selon qu'elles soient prononcées par les Tchèques de Bohême ou les Tchèques de Moravie (Simackova *et al.*, 2012). Dans la région de Bohême, les locuteurs distingueraient les voyelles de la paire *i/i:* essentiellement au niveau spectral. En revanche, les Tchèques de Moravie produiraient les deux voyelles avec des caractéristiques acoustiques plus proches mais avec une différence de durée plus marquée. Le triangle vocalique du tchèque de Bohême (à gauche) et du tchèque de Moravie (à droite), selon Simackova *et al.* (2012) est illustré à la Figure 16. Nous remarquons que le *[i]* bref, tel qu'il est prononcé dans la région de Moravie est plus extrême que le *[i]* bref prononcés par les Tchèques de Bohême.

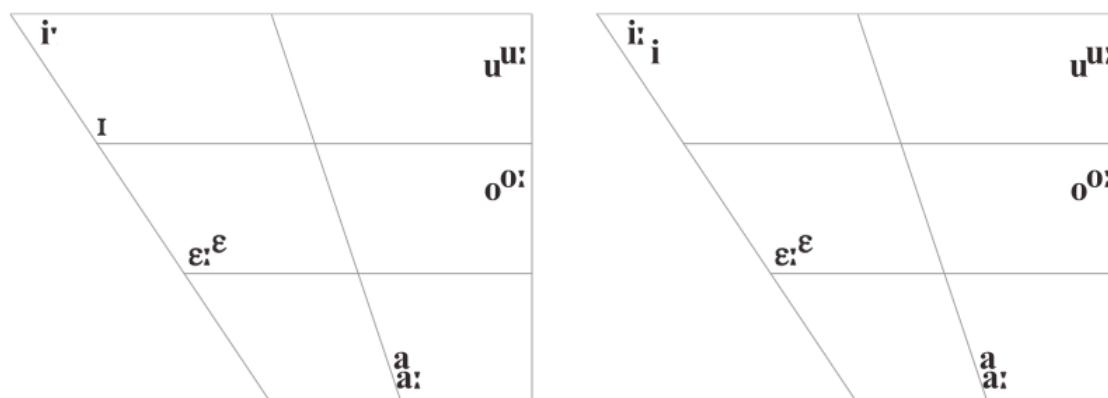


Figure 16 : Le triangle des voyelles monophthongues du tchèque de Bohême (à gauche) et de Moravie (à droite), selon Simackova *et al.* (2012)

Afin d'attester du caractère phonémique des dix voyelles monophthongues ainsi que de celui des trois diphtongues tchèques, nous utilisons la méthode de paires minimales (Troubetzkoy, 1957) qui se trouvent dans le Tableau 8.

/bi:t/ býť 'être'	/bit/ byt 'l'appartement'
/le:tu/ létu 'de l'été'	/letu/ letu 'du vol'
/ra:n/ rán 'de matins'	/ran/ ran 'de coups'
/bo:je/ bóje 'les bouées'	/boje/ boje 'les combats'
/domu:/ domů 'à la maison, chez nous'	/domu/ domu 'du bâtiment'
/eura/ eura 'les euros'	/aura/ aura 'l'aura'
/houba/ houba 'le champignon'	/huba/ huba 'la gueule'

Tableau 8 : Paires minimales attestant du caractère phonémique des voyelles du tchèque

Il est à noter que les diphtongues /eu/ et /au/ de même que la monophthongue longue /o:/ n'apparaissent que dans des mots d'emprunt.

2.2.1.2 Inventaire vocalique du français

En français, le nombre de voyelles, composant le système phonologique d'un locuteur, dépend de son arrière-plan diatopique (dialecte), diastratique (génération, appartenance socio-professionnelle) et diaphasique (registre) (Valdman, 2000). Le système vocalique maximale est composé de 15 voyelles, onze voyelles orales /i, e, ε, a, y, ø, œ, u, o, ɔ, α/, comme illustrées à la

Figure 17 (entourées), reprise de Vaissière (2009) et quatre nasales /ẽ, ẽ̃, ỹ, ỹ̃/.

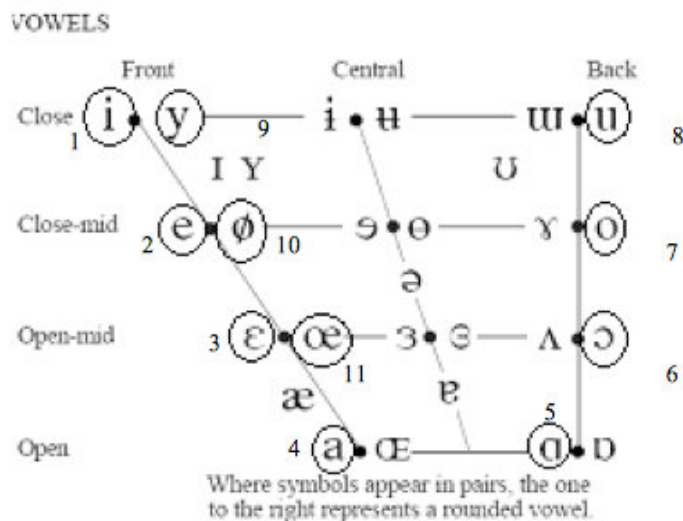


Figure 17 : 11 voyelles orales du français (entourées), selon Vaissière (2009)

L'inclusion de la voyelle centrale non labialisée /ə/ est incluse dans le système phonologique du français dépend des auteurs (pour revue de littérature, voir Fougeron *et al.* (2007). Certains chercheurs indiquent que le schwa se réalise en français [œ] (Blanchet, 1993; Féry, 2011) ou [ø] (Duez and Dubeda, 2006) alors que d'autres stipulent que sa réalisation acoustique diffère de ces deux sons vocaliques (Adda-Decker *et al.*, 1999; Fougeron *et al.*, 2007). Wioland (2005, p. 44) explique la raison pour laquelle /ə/ n'est pas inclus dans l'enseignement de FLE : « Nous ne pouvons pas présenter le

2. Le tchèque et le français

symbole /ə/ dans ce cadre, puis qu'il ne peut représenter la prononciation d'une voyelle en dernière syllabe prononcée de 'mot phonétique'. »

Dans le français actuel, les voyelles /a/ et /œ/ permettant opposer respectivement « âne » et « Anne », puis « brun » et « brin », disparaissent au profit de /a/ et /ɛ̃/ ce qui aboutit vers un système à 13 voyelles. Detey and Racine (2012) conseillent de présenter le système à 13 voyelles françaises aux apprenants de FLE et Wioland (2005, p. 43) ajoute que : « *il s'agit d'un nombre volontairement minimal de phonèmes dans le cadre d'un modèle didactique pour l'acquisition du français parlé tel qu'il est en usage de nos jours.* ».

Enfin selon Léon (1976), le système vocalique minimale ne comporte plus que dix phonèmes en français. En effet dû à la neutralisation des contrastes entre les voyelles moyennes e/ɛ, ø/œ et o/ɔ (qui se trouvent de plus en plus en distribution complémentaire en français actuel), les voyelles moyennes mi-fermées tendent à apparaître en syllabe ouverte et inversement les voyelles mi-ouvertes en syllabe fermée.

Notons que contrairement au tchèque, la durée vocalique n'a plus de fonction phonologique en français standard. La durée est cependant contrastive en final du mot dans de nombreux dialectes du français, comme en français neuchâtelois, opposant les voyelles brèves /i, y, e, u/ et longues /i:, y:, e:, u:/ (Dommergues, 2008; Grosjean *et al.*, 2007).

Les paires minimales pour attester des treize voyelles du français actuel sont indiquées dans le Tableau 9 (inspirées de Vaissière (2006) :

/pil/ pille	/pyl/ pull
/pul/ poule	/pol/ pôle
/pəl/ Paul	/pal/ pâle
/ʒœn/ jeûne	/ʒœn/ jeune
/le/ les	/lɛ/ laid
/vɛ̃/ vin	/vɑ̃/ vent
/vɑ̃/ vent	/vɔ̃/ vont

Tableau 9 : Paires minimales attestant du caractère phonémique des voyelles du français

Les contrastes vocaliques dans les deux langues reposent sur le paramètre de l'aperture, avec quatre degrés pour le français et trois degrés pour le tchèque, de l'antériorité/ postériorité, puis de la labialité et nasalité pour le français et de la durée pour le tchèque.

Le Tableau 10 compare les voyelles du tchèques et du français ainsi que leur distribution.

VOYELLES	En français	En tchèque
Plan distributionnel	peuvent occuper toutes les positions : initiale/ finale/ interconsonantique peuvent apparaître en contigu ex : aéroport	peuvent occuper toutes les positions : initiale/ finale/ interconsonantique peuvent apparaître en contigu ex : neutopil se
Début vocalique	- attaque douce	- attaque dure, insertion d'un coup de glotte
Diphthongues	- inexistantes	- /au/, /eu/, /ou/
Voyelles antérieures arrondies	- /y/, /ø/, /œ/	- inexistantes
Voyelles nasales	- /ẽ/, /ũ/, /ĩ/	- inexistantes
Voyelles longues	- inexistantes	- /i:/, /e:/ [ɛ:], /a:/, /u:/, /o:/
Nombre de degrés d'aperture	4, opposant voyelles mi-ouvertes et mi-fermées	3

Tableau 10 : Tableau comparatif des voyelles du tchèque et du français

2.2.1.3 Etudes acoustiques sur les voyelles

2.2.1.3.1 Voyelles du français

Les fréquences de résonance des voyelles du français traditionnellement citées sont celles de Delattre (1964) et de Calliope (1989), mais elles ne devraient plus être utilisées.

Les premier et deuxième formants des voyelles du français, proposés par Delattre (1964), sont établis à partir de paramètres utilisés dans la synthèse vocale. Ils sont calculés pour représenter les voyelles orales, produites par un locuteur masculin, avec la fréquence fondamentale à 120 Hz. Les données ne proviennent pas de l'analyse de corpora authentiques.

Calliope (1989) propose une autre série de valeurs formantiques pour les voyelles du français, en mentionnant que « *les données de Delattre, 1965, ne doivent plus être utilisées.* » Dans Calliope, les valeurs de F1, F2, F3 et F4 sont calculées à partir de productions de dix hommes et neuf femmes. Ainsi, les voyelles [e, o, u, y, ø] sont insérées en syllabe ouverte, précédées d'une occlusive labiale pV alors que les voyelles [i, ɛ, a, ɔ, œ] sont produites en syllabe fermée, précédées d'une occlusive labiale et suivies d'une fricative uvulaire pVR. La disparité des contextes et le contexte uvulaire élevant la fréquence du premier formant vocalique et abaissant la fréquence du deuxième formant rend les résultats de cette étude moins pertinents.

Les études plus récentes sur les formants des voyelles du français ont été menées par exemple par Gendrot and Adda-Decker (2005) dans des grands corpora ou Georgetown *et al.* (2012) pour les voyelles isolées. Les données de Gendrot and Adda-Decker (2005), issues de la parole radiophonique (deux heures d'enregistrement d'émissions), mettent en évidence les valeurs des trois premiers formants des dix voyelles orales [i, e, ɛ, a, u, o, ɔ, y, ø, œ]. Les valeurs moyennes sont calculées séparément à partir de productions de 15 femmes et 15 hommes. Le Groupe Didactique (GD) du LPP

2. Le tchèque et le français

de Paris 3 a établi les valeurs formantiques des quatre premiers formants à partir de la production hautement contrôlée de voyelles isolées, répétées quatre fois, par 40 femmes (Georgeton *et al.*, 2012). Les valeurs de Gendrot et Adda-Decker (formants des voyelles produites par les femmes uniquement, tout contexte consonantique confondu) et de Georgeton *et al.* (2012) sont indiquées dans le Tableau 11 et illustrés à la Figure 18. Les écarts types se trouvent entre parenthèses.

	Moyenne de F1		Moyenne de F2		Moyenne de F3		Moyenne de F4	
	GD	G&A	GD	G&A	GD	G&A	GD	G&A
i	275 (32)	348 N/D	2585 (228)	2365 N/D	3815 (228)	3130 N/D	4521 (256)	N/D
e	405 (44)	423 N/D	2553 (174)	2176 N/D	3346 (202)	2860 N/D	4325 (271)	N/D
ε	614 (83)	526 N/D	2306 (160)	2016 N/D	3137 (202)	2800 N/D	4383 (271)	N/D
a	830 (113)	685 N/D	1438 (183)	1677 N/D	2900 (179)	2735 N/D	4065 (256)	N/D
y	276 (29)	371 N/D	2091 (167)	2063 N/D	2579 (216)	2745 N/D	3826 (221)	N/D
ø	409 (47)	420 N/D	1599 (162)	1693 N/D	2703 (178)	2687 N/D	3985 (190)	N/D
œ	599 (86)	436 N/D	1678 (156)	1643 N/D	2843 (208)	2715 N/D	4107 (221)	N/D
u	291 (31)	404 N/D	779 (93)	1153 N/D	2648 (254)	2742 N/D	3980 (356)	N/D
o	415 (44)	438 N/D	842 (103)	1140 N/D	2862 (165)	2790 N/D	4048 (228)	N/D
ɔ	595 (100)	528 N/D	1144 (141)	1347 N/D	2907 (172)	2743 N/D	4035 (209)	N/D

Tableau 11 : Valeurs moyennes des formants F1, F2, F3, F4 pour les dix voyelles orales du français [i, e, ε, a, u, o, ɔ, y, ø, œ], selon Georgeton *et al.* (2012)(GD), calculées à partir de 1600 voyelles isolées (10 voyelles*4 répétitions*40 locutrices) et Gendrot et Adda-Decker (2005) (G&A), calculées à partir de 2 heures de parole radiophonique (15 femmes, tout contexte consonantique). Les écarts types sont entre parenthèses (avec les valeurs extrêmes grisées), N/D : non défini.

Les écarts-types indiqués nous renseignent sur la variabilité inter et intra-locuteur et ils varient en fonction de la voyelle. Les voyelles fermées, qui ont un F1 bas, présentent en général une valeur moyenne du premier formant avec un écart type plus faible que les voyelles ouvertes, qui ont un F1

élevé. Ainsi, la moyenne et l'écart type du F1 indiqué par le GD est de 275 Hz (32 Hz) pour la voyelle [i] (variation de 11,6 %), et de 830 Hz (113 Hz) pour la voyelle [a] (variation de 13,6 %).

La Figure 18 représente le triangle vocalique F1/F2 (en Bark) des dix voyelles orales du français pour des voix féminines selon les données de Gendrot et Adda-Decker (trait pointillé) et de Georgetown *et al.* (2012) (trait plein). Les valeurs formantiques reflètent le type de corpus utilisé. Ainsi, en parole spontanée, Gendrot et Adda-Decker relèvent des valeurs de formants plus centralisées par rapport aux valeurs mesurées à partir de corpora contrôlés. Les voyelles prononcées en isolation, étudiées par Georgetown *et al.* (2012) fournissent des valeurs formantiques extrêmes, créant ainsi un triangle vocalique maximal. Les voyelles [i, y, u, e, ø, o] se réalisent avec un premier formant plus bas que les voyelles [ɛ, œ, ɔ, a]. Les voyelles antérieures [i, y, e, ɛ] présentent un deuxième formant plus élevé que les voyelles postérieures [u, o, ɔ, a].

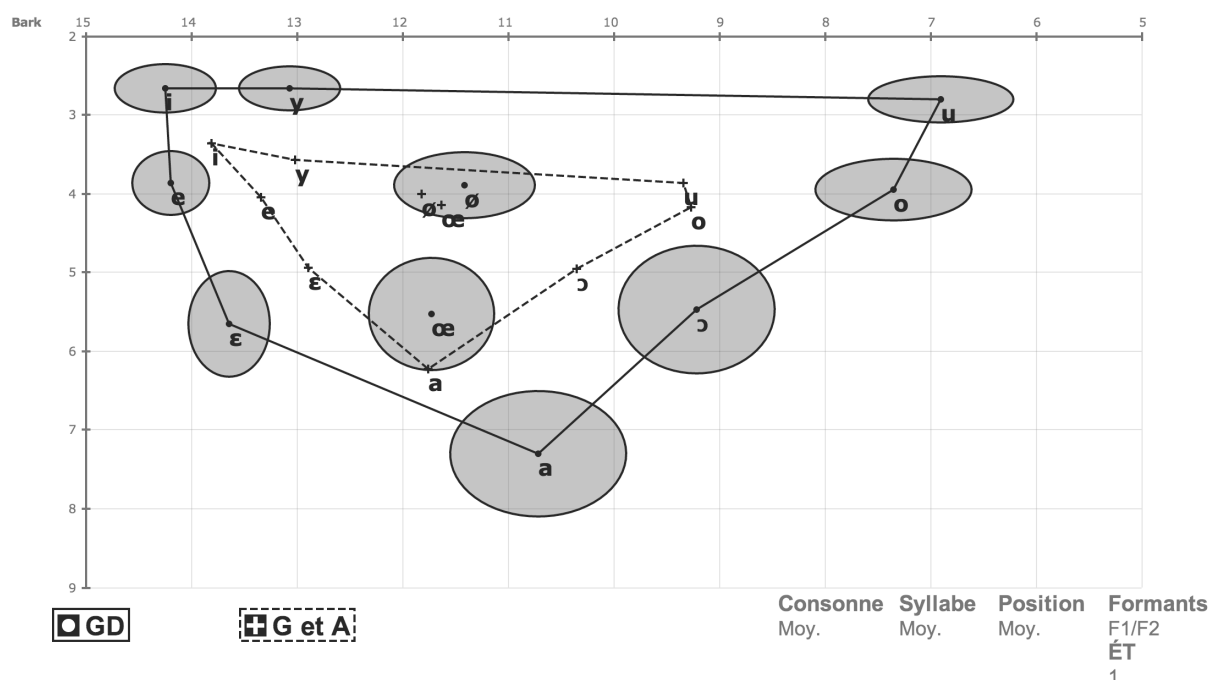
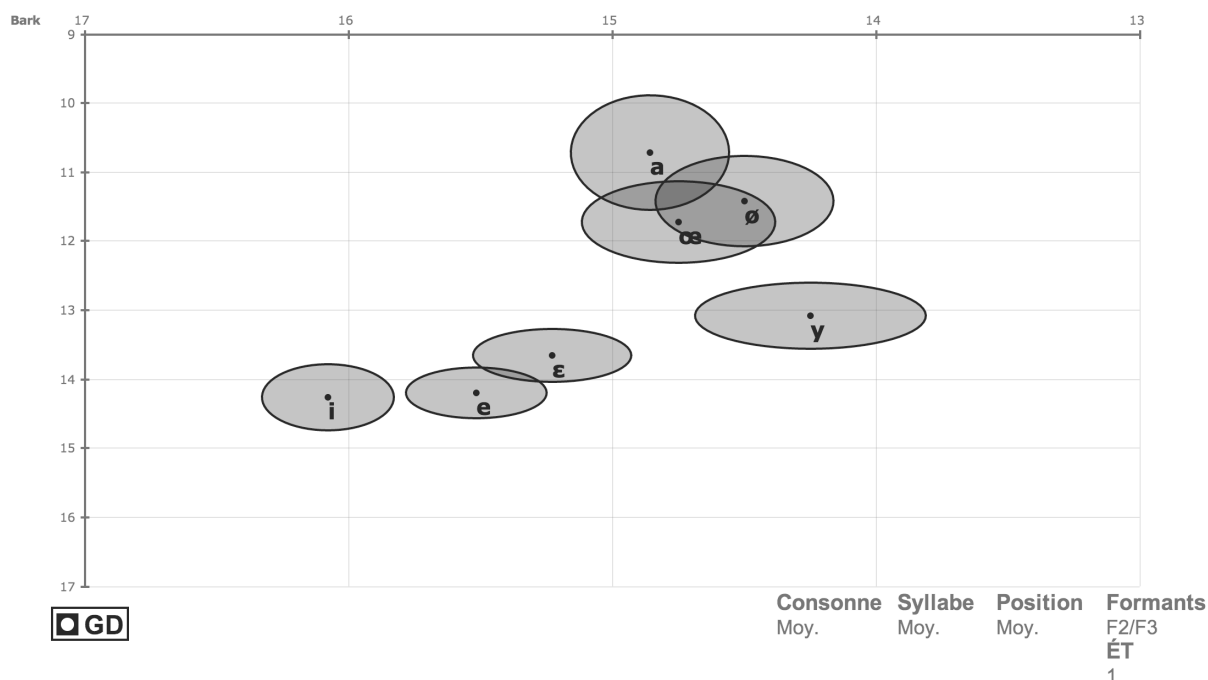


Figure 18 : Comparaison des triangles vocaliques sur le plan F1-F2 (en Bark) à partir de données de Gendrot and Adda-Decker (2005) : moyennes de productions de 15 femmes, 2 heures de parole radiophonique (en pointillé), et du Groupe Didactique (Georgetown *et al.*, 2012) : moyennes de productions de 40 femmes, 4 répétitions, voyelles isolées (en trait plein). Les ellipses sont tracées à un écart type des moyennes F1-F2

La Figure 19 représente le triangle vocalique sur le plan F2/F3 (en Bark), calculé à partir de valeurs formantiques de Georgetown *et al.* (2012) pour les voyelles isolées. Les ellipses sont tracées à un écart type de la moyenne.

2. Le tchèue et le français



La Figure 19 montre que la voyelle labiale [y] se réalise avec le F3 minimal alors que [i] étiré avec le F3 maximal. Puis [ø] et [œ] se réalisent avec F3 plus bas que [e] et [ε], comme attendu.

2.2.1.3.2 Voyelles du tchèue

Le caractère acoustique des voyelles du tchèue est originellement décrit dans les travaux de Hála (1941) qui a mesuré les formants vocaliques F1 et F2 à partir de la lecture de mots, de phrases et de la parole spontanée par quatre locuteurs natifs. Les valeurs moyennes (en Hz) sont illustrées dans le triangle vocalique F1/F2 de la Figure 20. Nous y remarquons que outre la paire e:/e, les voyelles longues sont définies par des valeurs plus extrêmes que leurs homologues brèves. Ainsi, les voyelles arrière longues [u:] et [o:] sont plus fermées et plus postérieures que [u] et [o] brefs respectivement, le [a:] long est plus ouvert et plus postérieur que le [a] bref et la voyelle longue [i:] est plus fermée et plus antérieure que [i] bref.

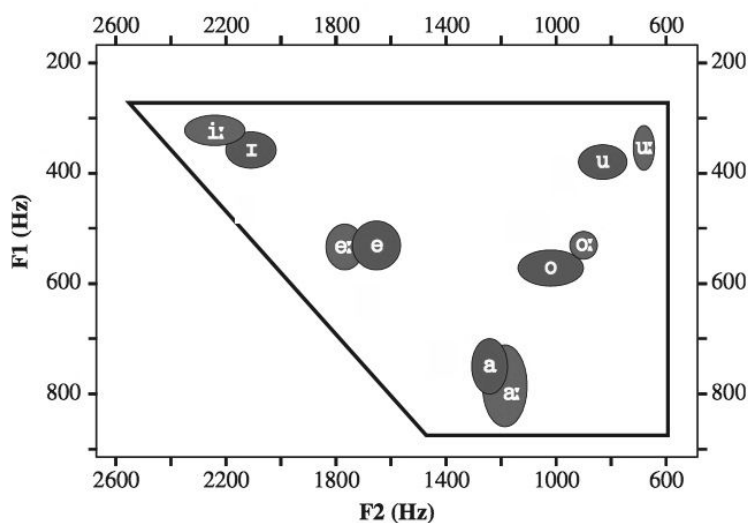


Figure 20 : Triangle vocalique F1/F2 (en Hz) des voyelles tchèques à partir de productions de 4 locuteurs natifs, selon Hála (1941)

Les descriptions plus récentes, basées sur un grand nombre de locuteurs, sont celles de Grepl *et al.* (2007) ou Skarnitzl and Volin (2012). Les premiers fournissent des zones de valeurs des premier et deuxième formants, chaque zone étant définie par les valeurs minimale et maximale du formant. Les valeurs sont issues de productions de 35 locuteurs (21 hommes et 14 femmes) de voyelles brèves isolées produites en voix chuchotée. Leur position dans le triangle vocalique F1/F2 est illustrée à la Figure 21.

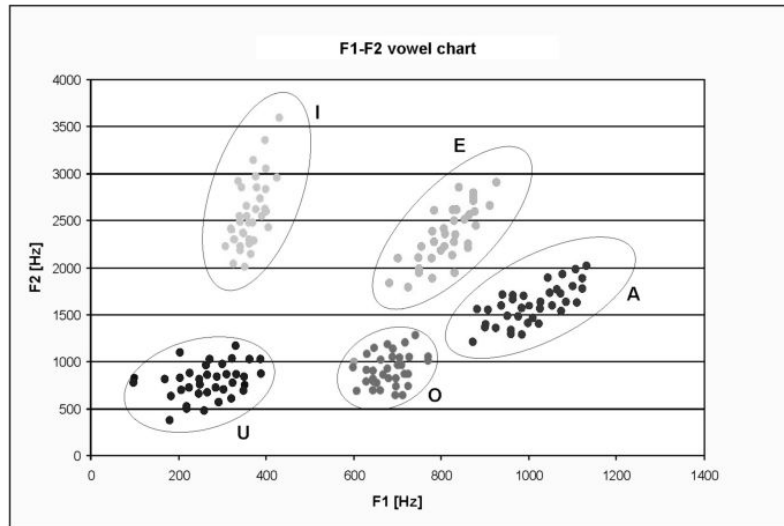


Figure 21 : Triangle F1/F2 (en Hertz) des voyelles brèves du tchèque [ɪ, ε, u, a, o] produites en isolation par 21 hommes et 14 femmes, selon Grepl *et al.* (2007)

Skarnitzl and Volin (2012) s'intéressent aux valeurs des premier et deuxième formants des voyelles brèves et longues, calculées à partir de la tâche de lecture par 27 hommes et 48 femmes séparément. La Figure 22 représente le triangle vocalique F1/F2 (en ERB-Hz) avec les valeurs moyennes calculées à partir de productions de 27 hommes (à gauche) et de 48 femmes (à droite). L'écart type affiché est de un.

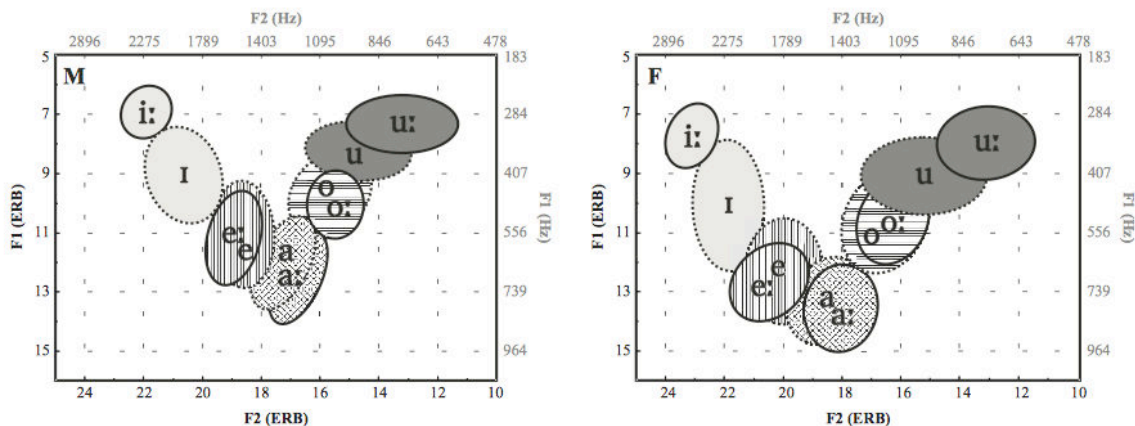


Figure 22 : Triangle F1/F2 (en ERB-Hz) des voyelles du tchèque [ɪ, i:, ε, ε:, a, a:, o, o:, u, u:] issue de la tâche de lecture par 27 hommes (à gauche) et 48 femmes (à droite), selon Skarnitzl and Volin (2012)

Le Tableau 12 permet de comparer les valeurs de Grepl *et al.* (zones formantiques définies par la valeur minimale et maximale pour les voyelles brèves uniquement) et celles de Skarnitzl et Volin (valeurs moyennes avec les écarts types calculées à partir de productions de 27 hommes, pour les voyelles brèves et longues).

2. Le tchèque et le français

	(Zone de) F1		(Zone de) F2	
	Grepl et al.	S&V	Grepl et al.	S&V
ɪ	300-450	415 (77)	2000-3600	1943 (202)
i:	N/D	282 (33)	N/D	2255 (152)
ɛ	700-950	566 (99)	1700-3000	1519 (128)
ɛ:	N/D	576 (86)	N/D	1578 (123)
o	600-800	458 (69)	600-1400	1055 (130)
o:	N/D	484 (55)	N/D	1028 (82)
u	100-400	359 (43)	400-1200	937 (145)
u:	N/D	304 (38)	N/D	769 (132)
a	850-1150	649 (95)	1200-2000	128 (114)
a:	N/D	699 (97)	N/D	1206 (97)

Tableau 12 : Les zones de formants F1 et F2 pour les voyelles brèves, selon Grepl et al., et les moyennes formantiques (avec les écarts types entre parenthèses) pour les voyelles brèves et longues du tchèque de 27 hommes, selon Skarnitzl et Volin (S&V). N/D signifie non défini.

Les résultats des deux études diffèrent et la différence peut être expliquée par le choix du corpus. En lecture et donc dans des contextes consonantiques divers, les voyelles ouvertes [ɛ, ɔ, a] rencontrent des valeurs de F1 plus bas que lorsque les voyelles sont prononcées en isolation alors que le F1 des voyelles fermées [ɪ, u] en parole lue se rapproche de la limite supérieure de la zone formantique F1 indiquée pour les voyelles fermées produites en isolation. En termes de F2, les voyelles antérieures [ɪ, ɛ] issues de la lecture rencontrent des valeurs plus basses que celles des zones de formant F2 calculées à partir des mêmes voyelles produites à l'isolée. Ceci indique une réduction des voyelles produites dans des contextes consonantiques divers (en parole lue) par rapport à la production de voyelles isolées.

Ainsi contrairement à Grepl et al., Skarnitzl et Volin étudient les valeurs des formants des voyelles brèves et également des voyelles longues. Alors qu'il est généralement admis qu'en tchèque actuel, les voyelles de la paire *i/i:* ont un timbre différent, les auteurs suggèrent une différence de timbre également au niveau de la paire *u/u:*. En effet, après avoir transformé les valeurs en unités ERB, Skarnitzl et Volin notent une différence de 12,6 % dans le cas du F1 de *u:/u* et de 11,4 % dans le cas de leur F2, comme indiqué dans le Tableau 13. Ils ont alors conclu qu'en relation avec le JND des formants fixé à 5 %, la différence des formants du [u] bref et du [u:] long devrait être phonétiquement pertinente et conduire vers la perception de deux timbres distincts. Notons néanmoins que ce postulat n'a pas été vérifié dans des tests de perception.

	F1 (Hz)	F2 (Hz)	F1 (ERB)	F2 (ERB)	rozdl F1 %	rozdl F2 %
i	414,9 (76,6)	1943,2 (202,0)	9,06 (1,08)	20,64 (0,87)	30,55	5,80
i:	281,7 (33,2)	2255,2 (151,6)	6,94 (0,58)	21,91 (0,56)		
e	566,3 (98,5)	1518,8 (127,7)	11,05 (1,18)	18,62 (0,70)	1,16	1,69
e:	575,7 (86,0)	1578,0 (122,9)	11,18 (1,03)	18,94 (0,63)		
a	648,5 (94,8)	1286,0 (113,5)	12,01 (1,03)	17,26 (0,72)	4,23	3,05
a:	698,6 (96,5)	1206,3 (97,2)	12,54 (1,01)	16,75 (0,65)		
o	457,7 (69,3)	1054,8 (129,8)	9,68 (0,93)	15,66 (0,92)	3,68	1,16
o:	483,7 (54,9)	1027,9 (82,0)	10,05 (0,75)	15,48 (0,62)		
u	359,2 (43,1)	936,6 (145,1)	8,25 (0,66)	14,71 (1,19)	12,55	11,36
u:	304,1 (37,7)	768,9 (132,4)	7,33 (0,64)	13,21 (1,24)		

Tableau 13 : Les valeurs formantiques des voyelles tchèques, en Hertz, en ERB et la différence en pourcentage entre les valeurs formantiques des voyelles brèves et longues au niveau de chaque paire. Les écarts types sont entre parenthèses. Selon Skarnitzl and Volin (2012)

2.2.2 Systèmes consonantiques

2.2.2.1 Inventaire consonantique du tchèque

Le système consonantique du tchèque compte 25 phonèmes, qui sont représentées selon le mode et le lieu d'articulation dans le Tableau 14. Les phonèmes qui ne font pas partie du système consonantique du français y sont marqués en gras et il s'agit de /c, ʃ, ts, tʃ, x, h, r, ř/.

Les paires minimales permettant d'établir le caractère phonémique des consonnes du tchèque se trouvent dans le Tableau 15.

Il est à noter que le phonème /g/ ne figure que dans les mots d'emprunt et le phonème /f/ est attesté par seulement deux paires minimales composés de mots tchèques (dans les autres paires minimales, /f/ apparaît dans des mots d'emprunt).

Nous remarquons que le tchèque a deux consonnes vibrantes, le /r/, produit avec un à trois battements de l'apex contre les alvéoles, et le /ř/, produit avec plusieurs battements de moindre amplitude (Palková, 1997).

2. Le tchèque et le français

		bilabiale	labio-dentale	dentale/ alvéolaire	post- alvéolaire	palat ale	vélaire	glottale
Occlusive		p b		t d		c ʃ	k g	
Fricative			f v	s z	ʃ ʒ		x	h
Fricative vibrante				ř ⁵				
Affriquée				ts	tʃ			
Sonante	nasale	m		n		ɲ		
	approx.					j		
	latérale			l				
	vibrante			r				

Tableau 14 : Système consonantique du tchèque ; les phonèmes inexistant dans le système français sont en gras

/pa:t/ pád 'la chute'	/ba:t/ bát (se) 'avoir peur'
/ta:t/ tát 'fondre'	/da:t/ dát 'donner'
/celo/ tělo 'le corps'	/jelo/ dělo 'le canon'
/krok/ krok 'le pas'	/grog/ grog 'le grog'
/oni/ oni 'ils'	/oni/ ony 'elles'
/zouvat/ zouvat 'déchausser'	/zoufat/ zoufat 'désespérer'
/sa:ska/ sázka 'le pari'	/la:ska/ láska 'l'amour'
/zup/ zub 'la dent'	/dup/ dub 'le chêne'
/uʃi:t/ ušít 'coudre'	/uʒi:t/ užít 'profiter'
/xoʃit/ chodit 'marcher'	/hoʃit/ hodit 'lancer'
/tsop/ cop 'la tresse'	/mop/ mop 'le balai'
/tʃelo/ čelo 'le front'	/jelo/ jelo 'a roulé'
/hra:t/ hrát 'jouer'	/hřa:t/ hřát 'chauffer'

Tableau 15 : Paires minimales attestant du caractère phonémique des consonnes du tchèque

⁵ Notons que l'API ne possède pas de symbole pour la consonne /ř/, n'existant que dans la langue tchèque

2.2.2.2 Inventaire consonantique du français

Le nombre de consonnes en français ne varie pas d'une variété régionale à l'autre (Vaissière, 2006, p. 12). Ainsi, le système consonantique est composé de 17 consonnes /p, b, t, d, k, g, m, n, ɲ, f, v, s, z, ʃ, ʒ, ʁ, l/, classées selon le mode et le lieu d'articulation dans le Tableau 16 et de trois semi-consonnes /j, ɥ, w/. Nous remarquons que seulement un phonème ne fait pas partie du système tchèque : la fricative uvulaire /ʁ/ (marquée en gras). Les paires minimales qui attestent des consonnes du français se trouvent dans le Tableau 17 (liste reprise de (Vaissière, 2006)).

		bilabiale	labio-dentale	dentale/alvéolaire	post-alvéolaire	palatale	vélaire	uvulaire
Occlusive		p b		t d			k g	
Fricative			f v	s z	ʃ ʒ			ʁ
Sonante	nasale	m		n		ɲ		
	latérale			l				

Tableau 16 : Système consonantique du français ; les phonèmes qui ne font pas partie du système consonantique tchèque sont en gras

/pã/ pan	/bã/ banc
/fã/ faon	/vã/ vent
/tã/ temps	/dã/ dent
/kã/ camps	/gã/ gant
/sã/ sang	/zã/ zan
/ʃã/ chant	/ʒã/ gens
/mã/ ment	/lã/ lent
/lã/ lent	/ʁã/ rang
/aɲo/ agneau	/ano/ anneau

Tableau 17 : Paires minimales attestant du caractère phonémique des consonnes du français

Les paires minimales attestant des semi-consonnes du français se trouvent dans le Tableau 18.

/fiɲ/ fille	/fil/ file
/lwi/ Louis	/lɥi/ lui

Tableau 18 : Paires minimales attestant du caractère phonémique des semi-consonnes du français

Le Tableau 19 compare les consonnes du tchèque et du français, de même que leur distribution et combinaison.

CONSONNES	En français	En tchèque
Plan distributionnel	peuvent occuper toutes les positions : initiale/ finale/ intervocalique les consonnes sonores ne se dévoient pas en finale assimilation de voisement essentiellement régressive ⁶	peuvent occuper toutes les positions : initiale/ finale/ intervocalique les consonnes sonores se dévoient en finale absolue assimilation de voisement essentiellement régressive ⁷
Phonotactique	au maximum 3 consonnes en attaque, 4 en coda /l, r, m, n/ ne peut pas constituer le noyau de la syllabe	au maximum 4 consonnes en attaque, 3 en coda Le noyau syllabique peut être formé par /l, r, m, n/
Consonnes affriquées	- inexistantes	- /ts/, /tʃ/
Consonnes vibrantes	- inexistantes (en français standard)	- /r/, /ř/
Consonne glottale	- inexistante	- /h/
Consonne uvulaire	- ʁ	- inexistante

Tableau 19 : Tableau comparatif des consonnes du tchèque et du français

Le Tableau 19 montre qu'en finale du mot, les consonnes sonores sont systématiquement dévoisées en tchèque alors qu'elles gardent le trait sonore en français. Les règles phonotactiques du tchèque permettent un regroupement maximal de quatre consonnes en attaque, comme dans le mot *pstruh* /pstrux/ 'la truite', un noyau syllabique constitué d'une voyelle monophthongue ou diphtongue de même que d'une des quatre sonantes /l, m, n, r/, comme dans les mots *plž* /plʃ/ 'la limace', *Keaton* /kitn/ prénom 'Keaton', *sedm* /sedm/ 'sept', *krk* /krk/ 'le cou', et enfin au maximum trois consonnes en coda, comme dans *pomst* /pomst/ 'des vengeances'. Le français, en revanche, permet une suite de trois consonnes constituant l'attaque d'une syllabe, comme dans *stridant* /stRidã/, le noyau syllabique repose toujours sur une voyelle, orale ou nasale, et enfin le coda peut être composé de quatre consonnes au maximum, selon la base de données lexicales et syllabiques ULSID (Rousset, 2004).

2.3 Systèmes supra-segmentaux

La prosodie englobe les procédés d'accentuation lexicale, d'intonation et de rythme qui sont fortement liés et ne peuvent pas être clairement séparés l'un de l'autre sur le plan perceptif, comme ne

⁶ L'assimilation progressive en français se produit plus rarement et on la trouve par exemple dans le mot 'chevaux' [ʃvø]

⁷ L'assimilation progressive se produit lors que la vibrante /ř/ est précédée d'une consonne sourde, comme dans le mot *křik* [křik] 'le cri' (Dvorak, V. (2010). Voicing assimilation in Czech. In: *Rutgers Working Papers in Linguistics*, P. Staroverov, Ed, Vol. 3, Rutgers University, NJ : LGSA, pp. 115-144.

peuvent l'être les aspects segmentaux et les aspects prosodiques de la parole. Les mêmes corrélats acoustiques sont utilisés et correspondent aux variations de la fréquence fondamentale, de la durée, de l'intensité physique, de la qualité de la voix et aux variations allophoniques (Vaissière, 2006).

Fonagy (1980) définit l'accent du point de vue perceptif comme une mise en relief d'une syllabe par rapport aux autres syllabes de l'unité accentuelle. L'accent primaire en français, issu de l'accent lexical en latin, est fixe et il se trouve en finale de mot. Dans la parole continue, toutes les dernières syllabes de mots ne sont pas accentuées, seule la syllabe finale d'un groupe de sens (ou encore groupe rythmique) l'est clairement. Ainsi comme l'a rappelé Delattre (1939, p. 141), « *l'accent appartient non au mot mais au groupe [...]. Ces groupes correspondent généralement aux groupes de sens, exprimant une unité sémantique.* ». Vaissière (2010) dénomme le français alors comme une « langue à frontières ». Le corrélat principal de l'accent primaire (lexical) de mot en français est la durée. Selon Delattre (1939, p. 142) : « *C'est celui des trois éléments acoustiques qui est le plus étroitement uni à l'accent et c'est le seul des trois qui en soit toujours un facteur.* ». Au niveau du groupe de souffle, entre deux pauses respiratoires, on observe une récurrence de deux mouvements mélodiques qui se terminent par une syllabe soit nettement montante, après un retour sur la ligne de base, soit par deux ou trois syllabes nettement descendantes (Vaissière, 2001). Notons que les montées mélodiques, en fin de groupe, ont une forte saillance perceptive en français.

Outre l'accent primaire qui est final, il existe en français actuel une tendance à renforcer également le début du mot (Vaissière, 2001). Cet accent en début de mot est dit initial et il a une origine emphatique (Garde, 1968). L'accent initial se manifeste par une forte pression sous-glottique. La syllabe sur laquelle le locuteur insiste est alors réalisée comme plus intense, avec un fondamental élevé, les consonnes occlusives en début de mot peuvent être aspirées ou allongées, les fricatives plus fermées et les voyelles précédées d'un coup de glotte. Ainsi, selon Vaissière (2001, p. 16) « *allongement, aspiration, coup de glotte et ton Haut sont donc liés sur le plan physiologique et hautement compatibles.* ». Notons que la première syllabe peut aussi au contraire perdre toute prééminence perceptive qui se déplace alors sur la seconde syllabe, notamment quand elle comporte un *e muet*, comme dans le mot *recompter* ou quand les mots commencent par une voyelle comme dans *épouvantable* (Fonagy, 1980).

Le tchèque fait également partie des langues à accent fixe. Contrairement au français, l'accent primaire est placé sur la première syllabe d'une mesure rythmique mis à part dans les mots d'origine étrangère, comme dans *ahoj* [aɦoj] 'salut' (origine anglaise) ou *pardon* [pardon] (origine française), qui peuvent être prononcés avec un accent initial ou final (Dubeda, 2002). Lorsqu'un mot lexical est précédé ou suivi d'un mot grammatical, il peut former avec ce dernier une seule mesure rythmique et l'accent se réalise alors à son début (Rigault, 1970).

Les corrélats acoustiques de l'accent en tchèque sont faibles (Volin and Studenovsky, 2007). Traditionnellement, on stipule que la prééminence accentuelle est créée par l'ajout de l'intensité (Hála, 1941) mais les études instrumentales plus récentes montrent que l'intensité ne joue qu'un rôle secondaire et que l'accent se manifeste essentiellement par les variations complexes de la fréquence fondamentale (Palkova, 1997). La durée, étant utilisée pour créer des contrastes phonologiques entre les voyelles brèves et leurs homologues longues, n'est pas classée parmi les corrélats accentuels malgré l'allongement final qui est attesté en tchèque (Dankovicova, 1997b).

L'accent crée ensuite une prééminence rythmique. Fraisse (1974) décrit deux types de rythme chez les humains : le rythme intensif où la syllabe la plus intense est souvent interprétée comme une syllabe initiale (comme en tchèque), et le rythme temporel où la syllabe allongée est interprétée comme une syllabe finale (comme en français).

Au niveau de la syllabe qui est la plus petite unité rythmique, il existe selon les langues une tendance ouvrante (répétition de syllabes ouvertes), comme en tchèque ou davantage encore en français et en italien, ou fermante (répétition de syllabes fermées), comme en anglais ou en allemand.

2. Le tchèque et le français

La Figure 23 illustre la répartition des structures syllabiques en allemand, anglais, tchèque, français et italien, selon Dankovicova and Dellwo (2007).

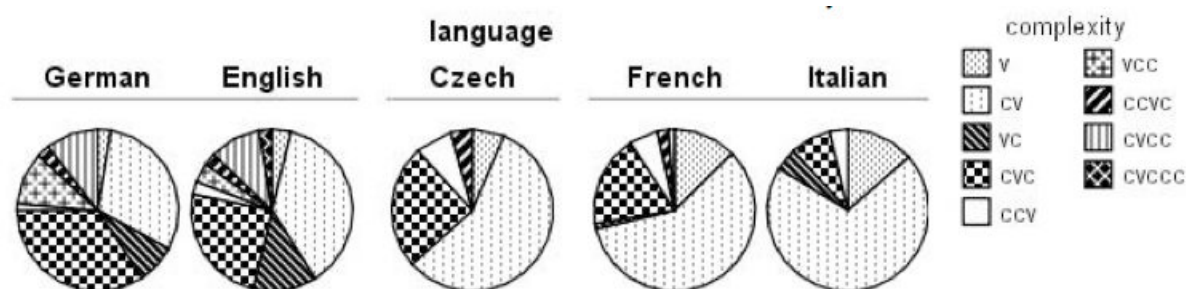


Figure 23 : Complexité des structures syllabiques en allemand, anglais, tchèque, français et italien, selon Dankovicova and Dellwo (2007)

Ainsi le système français possède environ 80 % de syllabes ouvertes (Wioland, 1991) et le tchèque, si on en croit Cechova *et al.* (2000) comporte deux fois plus de syllabes ouvertes que de syllabes fermées.

Le rythme du français et du tchèque est traditionnellement comparé à la mitrailleuse avec une isochronie syllabique (langue *syllable-timed*). L'isosyllabité des deux langues a été cependant remise en question par Wenk and Wioland (1982) pour le français et par Dubeda (2004) ou Dankovicova and Dellwo (2007) pour le tchèque. En effet, selon les descripteurs de classification rythmique, le tchèque posséderait les deux rythmes « isosyllabique » et « isochronique ». Dubeda (2012, p. 51) conclue cette problématique en disant que : « *Malgré cela, il n'est peut-être pas erroné de dire que, toutes analyses confondues, on a affaire à une langue plutôt 'isosyllabique'.* ».

3. Apprentissage phonétique

If you talk to a man in a language he understands, that goes to his head. If you talk to him in his language, that goes to his heart.

Nelson Mandela

Résumé : Dans le **troisième chapitre**, nous nous intéressons à l'apprentissage phonétique, qui concerne l'aspect supra-segmental (apprentissage du rythme, de l'accentuation, de l'intonation et de la qualité de la voix) et l'aspect segmental (voyelles et consonnes, leur combinaison, leur distribution en fonction de la structure syllabique et leur coarticulation). Les objectifs phonétiques dans l'apprentissage du Français Langue Etrangère (FLE) ont été largement développés par Delattre dont nous résumons les travaux. L'auteur insiste sur les particularités du français correspondant à l'habitude de tension, de croissance et d'antériorité.

En général, la difficulté accrue dans l'apprentissage des voyelles par rapport à celui des consonnes réside dans des gestes articulatoires moins bien définissables, dans des phénomènes de compensation (qui témoignent de la primauté de la cible acoustico-perceptive (Maeda, 1989a) ainsi que dans leur élasticité (leur nombre peut varier suivant la région d'un pays contrairement aux consonnes (Duez, 2001).

Nous rappelons également quelques études antérieures sur l'apprentissage des voyelles du français et les modèles qui permettent de prédire le niveau de difficulté. Alors que la théorie de l'interférence et l'analyse contrastive (Lado, 1957) sont basées sur des unités abstraites, en définissant les phonèmes *similaires* comme faciles à acquérir et les phonèmes nouveaux comme plus difficiles, les modèles récents tels que le Speech Learning Model « SLM » (Flege, 1995), le Perceptual Assimilation Model « PAM » (Best, 1995) ou le Native Language Magnet « NLM » (Kuhl and Iverson, 1995) opèrent avec des unités concrètes – les sons. Il s'agit essentiellement de modèles de perception qui montrent l'influence de la LM sur l'apprentissage de la LE. Alors que SLM et NLM sont des modèles acoustiques qui stipulent la nécessité de passer par des représentations mentales pour interpréter le signal acoustique, PAM est un modèle articulatoire direct qui considère que le geste articulatoire véhicule directement le sens linguistique. SLM et NLM définissent les sons nouveaux comme plus faciles à acquérir authentiquement que les sons *similaires* aux sons maternels. PAM s'intéresse aux contrastes en LE : le plus facile à percevoir est celui où les deux sons de la LE renvoient à deux sons différents de la LM. SLM, contrairement aux PAM et NLM, admet que toutes les erreurs en production ne s'expliquent pas par une perception erronée. Il existe en effet des cas où l'apprenant produit un son de manière authentique mais il le perçoit avec difficulté.

En didactique des langues, l'apprentissage est décrit comme une activité mentale qui induit une progression entre le début et la fin de la formation (Billières and Spanghero-Gaillard, 2005), laquelle se déroule donc par étapes. Chaque étape intermédiaire définit l'interlangue momentanée de l'apprenant (Bent and Bradlow, 2003; Flege, 1980; Fourakis and Iverson, 1985) qui se rapproche de la norme phonétique à petit pas. Cette évolution peut être observée grâce aux mesures acoustiques (Flege, 1980) mais tant que l'apprenant n'a pas atteint la cible acoustico-perceptive telle qu'elle est réalisée par un natif, le progrès est difficilement audible.

L'apprenant peut être défini comme une personne non-native qui apprend activement la langue cible, dans le but de communiquer (Best 2007).

Flege (1988b) définit la langue native (langue maternelle LM/langue première L1) comme étant la première langue apprise dans la petite enfance. Par opposition, la langue cible (langue étrangère LE/langue seconde L2) représente toute autre langue qui a été apprise après la LM, de

manière naturelle ou formelle. Par conséquent, un locuteur natif est la personne qui parle sa LM alors qu'un locuteur non-natif est celui qui parle une LE.

Les objectifs de l'apprentissage phonétique varient selon les approches. Pour leur part, (Harmegnies *et al.*, 2005, p. 273) définissent les objectifs par la négative : « *l'apprenant de la L2 a atteint les objectifs terminaux de la formation si, pour son interlocuteur natif, il s'exprime sans accent étranger.* ». Cet objectif peut néanmoins sembler trop ambitieux (Valdman, 2000), voire inatteignable si l'on en croit Ricci (1936, p. 364) : « *Le pur accent n'existe pas, prononcer une langue étrangère comme un enfant du pays, c'est une chimère.* ». Pour Flege and Hillenbrand (1984), il est peu probable qu'un apprenant tardif parvienne à produire certains types de sons de manière native. Par ailleurs, les approches récentes, notamment communicative ou actionnelle, fixent un but purement communicatif qui ignore la performance native par les non-spécialistes (Derwing and Munro, 2009; Leather, 1983).

3.1 Objets de l'apprentissage phonétique

Selon Wachs (2011), l'apprentissage phonétique requiert la capacité à oublier les motifs phonétiques utilisés en LM. Ainsi, l'apprenant doit acquérir la prosodie de la L2, c'est-à-dire son rythme, intonation et accentuation. Il doit également apprendre les nouveaux phonèmes, leur combinaison possible au sein de la syllabe ainsi que tous les allophones qui les réalisent.

Les difficultés des apprenants se situent alors aux niveaux suprasegmental et segmental (Anderson-Hsieh *et al.*, 1992; Moulton, 1962).

1. **Le niveau suprasegmental** nécessite la maîtrise de l'intonation, du rythme, de l'accentuation et de la qualité de la voix (Anderson-Hsieh *et al.*, 1992; Freland-Ricard, 1996; McCarthy, 1991; Métral, 1967). L'intonation est le mouvement mélodique de la parole, caractérisé par des variations de la hauteur (Delattre, 1960) et l'accent peut être défini comme la mise en valeur d'une syllabe au sein de l'unité accentuelle (Fonagy, 1980). Alors que l'intonation fait partie de toutes les langues du monde (aucune langue n'est monotone), il existe des langues sans accent (les langues mono-syllabiques comme le thaï) (Garde, 1968). Le rythme est ensuite défini comme toute répétition à intervalles suffisamment réguliers pour être perçus (Delattre, 1960). Enfin, la base articulatoire correspond selon (Vaissière, 2001, p. 18) à « *la façon 'moyenne' de positionner ses organes vocaux, articulatoires et phonatoires : une posture générale et une dynamique des mouvements.* ». Elle varie selon les langues et même selon les dialectes (Esling, 1987). Sur cette base, (Delattre, 1953; Delattre, 1964) oppose le français, avec un caractère antérieur à l'anglais, ayant un caractère postérieur.

Concernant les difficultés que peuvent rencontrer les apprenants tchèques de FLE à ce niveau, nous pouvons citer le travail de Dubeda (2009, 2012). L'auteur s'est intéressé à la réalisation de l'accent initial, fortement présent dans le français actuel, par des apprenants tchécophones. Du fait de la position initiale de l'accent en tchèque mais des réalisations tonales différentes et de l'absence de l'arc mélodique qui englobe plusieurs mots formant une unité syntaxico-prosodique, le transfert observé de la LM en LE est positif mais aussi négatif. Ainsi, Dubeda (2012, p. 86) définit l'accent initial français comme une « *arme à double tranchant* » pour les apprenants tchécophones de FLE.

2. **Le niveau segmental** nécessite la maîtrise de l'inventaire phonético-phonologique d'une langue donnée, c'est-à-dire des phonèmes qui sont les plus petites unités (abstraites) distinctives de la langue (Duchet, 1981) et de leurs allophones, de la phonotactique de la langue (combinaisons possibles des segments au sein de la syllabe) et de la stratégie de coarticulation. Il est admis que le système phonologique possède un nombre d'unités plus petit que le système phonétique (MacKain, 1982; MacKain *et al.*, 1981) et que, de cette manière, les langues se distinguent par leur systèmes phonologiques et phonétiques : alors que dans une langue, deux sons peuvent contraster au niveau phonologique, dans une autre langue, ces mêmes sons ne sont que des allophones d'un même phonème (MacKain *et al.*, 1981) voire n'apparaissent pas.

3. Apprentissage phonétique

Confronté à la relation qui existe entre un phonème et ses allophones, l'apprenant se trouve face à une alternative. Soit le système de sa LM possède deux phonèmes (comme par exemple le /u/ et le /ɔ/ du néerlandais) là où la LE n'en a qu'un (/u/ en espagnol), soit il possède un phonème (par exemple /ɾ/ en japonais) là où la LE en a deux (/l/ et /ɹ/ en anglais). La Figure 25 montre les deux types de relation.

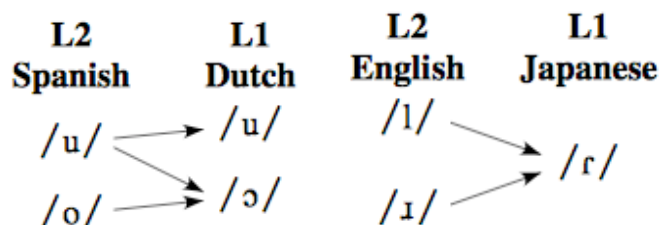


Figure 24 : Relation entre les phonèmes de la LM et de la LE: 2 catégories néerlandaises natives /u, ɔ/ correspondant à une catégorie espagnole non native /u/ (à gauche) ou une catégorie japonaise native /ɾ/ correspondant à 2 catégories anglaises non-natives /l, ɹ/ (split – scission - allophonique, à droite), selon (Escudero and Boersma, 2001)

Dans le premier cas de figure, l'apprenant peut percevoir les différents allophones de la L2 comme étant des phonèmes différents (Trofimovich *et al.*, 2001) et il doit réaliser que ces différents allophones correspondent à un seul phonème dans la L2 (Levy and Strange, 2008). Le deuxième cas de figure est appelé traditionnellement la « difficulté maximale » (Lado, 1957) ou plus récemment le « split allophonique »⁸ (Eckman *et al.*, 2001; Eckman *et al.*, 2003; Eckman and Iverson, 2013) car l'apprenant doit réaliser que les différents allophones d'un seul phonème de sa LM correspondent en effet à deux phonèmes distincts dans la LE. On évoquera à ce propos également le cas des Coréens qui apprennent les consonnes /s/ et /ʃ/ de l'anglais. En anglais, /s/ et /ʃ/ sont deux phonèmes distincts alors qu'en coréen, il s'agit de deux allophones du même phonème /s/. Les apprenants coréens doivent par conséquent réussir à scinder leur catégorie native en deux catégories non-natives.

Puisque l'apprenant doit faire face à ces nombreuses variantes, exprimant la même catégorie phonologique, les chercheurs (Flege, 1989; Flege, 1997; Levy, 2009) avancent que l'apprentissage phonétique se fait d'abord au niveau allophonique. Une fois que toutes les variations allophoniques sont maîtrisées, l'apprenant peut alors créer une nouvelle catégorie phonologique qui leur correspond. L'exemple des apprenants japonais ou chinois confirme le fait que l'apprentissage d'un son peut varier selon sa distribution. En effet, en chinois, les consonnes [t, d] peuvent contraster en position initiale mais jamais en position finale. Cette distribution a des conséquences pour les apprenants chinois de l'anglais qui n'arrivent pas à percevoir ce contraste en syllabe finale (Flege, 1989; Trofimovich *et al.*, 2001). Suivant la position, le contraste entre les consonnes provient d'indices acoustiques différents. Alors qu'en positions initiale et intervocalique, l'anglais se sert essentiellement du *Voice Onset Time* (VOT) pour différencier les consonnes occlusives sourdes et sonores, en position finale, c'est la durée de la voyelle précédente qui compte (MacKain, 1982). De même, les Japonais présentent de grosses difficultés à distinguer les consonnes /l/ et /ɹ/ de l'anglais en position initiale alors qu'en position finale la discrimination semble plus facile (Flege, 1995; Strange *et al.*, 2001; Trofimovich *et al.*, 2001).

De nombreux chercheurs s'accordent à dire que les personnes bilingues possèdent un espace phonologique commun pour les sons des deux langues qu'ils parlent (Best and Tyler, 2007) et afin de maintenir le contraste entre tous les sons au sein du système commun, les caractéristiques physiques de ces sons peuvent être modifiées par rapport à celles des locuteurs monolingues (Caramazza *et al.*, 1973). Fowler *et al.* (2008) examinent le VOT dans la production des consonnes occlusives [p, t, k] par des francophones et anglophones monolingues et bilingues de naissance et reportent qu'en français, les bilingues produisent des VOT significativement plus longs que les francophones monolingues et

⁸ Nous traduisons par « scission allophonique »

inversement, en anglais, les bilingues produisent des VOT significativement plus courts que les anglophones monolingues.

Concernant la combinaison des phonèmes, Hallé *et al.* (à paraître) reportent que la perception non-native est influencée par les règles phonotactiques de la LM. Ainsi, les Espagnols apprenant le français entendent et produisent le mot *spécial* en tant que [espesial] car la suite *sp est non attestée en espagnol. De même, les Français corrigent perceptivement la suite *tl, prononcée en début de mot en /kl/ selon les règles phonotactiques de leur LM (Hallé *et al.*, 1998).

Enfin, les stratégies de coarticulation diffèrent également selon les langues. Il s'agit d'une coordination de mouvements articulatoires résultant d'un long apprentissage qui est spécifique à chaque langue. Nous pouvons citer Grammont (1933, p. 21) à ce propos :

Cette coordination n'a été obtenue qu'après un long apprentissage. C'est à la suite de bien des essais plus ou moins heureux et de bien des tâtonnements que l'enfant arrive à prononcer exactement les sons de sa langue maternelle. Une fois qu'il a réussi et qu'il en a pris l'habitude, la coordination se fait d'elle même et inconsciemment.

3.1.1 Objectifs phonétiques en FLE (inspiré par les travaux de Pierre Delattre)

Cette section est inspirée par une série d'articles de P. Delattre, publiés dans la revue *The French Review* (Delattre, 1944; Delattre, 1945; Delattre, 1948; Delattre, 1960) et par son livret « *An Introduction to French Speech Habits* » (Delattre, 1947). Ces derniers sont consacrés à l'apprentissage phonétique d'une LE et notamment aux traits phoniques et prosodiques uniques au français. Pierre Delattre, qui a vécu entre 1903 - 1969, a effectué un travail d'époque dans le domaine de la phonétique instrumentale et expérimentale mais par dessus tout, comme le décrit son collègue de l'université d'Oklahoma Eddy (1974, p. 517) « *the full measure of the man : Pierre Delattre, teacher of French* ».

Selon Delattre (1948), trois habitudes articulatoires de base doivent être acquises lorsqu'on apprend le français : l'habitude de tension, d'antériorité et de croissance. Ces habitudes correspondent aux modes tendu, antérieur et croissant plus largement détaillés dans l'article de Delattre (1953).

D'abord, par l'habitude de tension des articulateurs et de croissance, l'apprenant doit aboutir à la maîtrise de :

- | | |
|----|---|
| 1. | Syllabation ouverte (prédominance de syllabes de type CV) |
| 2. | Egalité syllabique qui détermine en partie le rythme du français |
| 3. | Distribution d'intensité égale dans toutes les syllabes (pas de syllabes plus intenses que d'autres, mis à part en présence de l'accent emphatique) |
| 4. | Voyelles monophthongues (pas de diphtongaison) |
| 5. | Attaque vocalique douce, absence de coup de glotte |
| 6. | Non aspiration des occlusives sourdes (absence de surpression) |
| 7. | Non-affrication consonantique (mouvements des transitions formantiques pointus) |
| 8. | Détente des consonnes finales (avec l'ouverture de la bouche) |

Ensuite, par l'habitude d'antériorité l'apprenant du FLE doit aboutir à la maîtrise de :

- | | |
|----|--|
| 1. | Série de voyelles antérieures arrondies [y, ø, œ] |
| 2. | Consonnes [t̪, d̪, n̪, l̪, ʃ, ʒ] réalisées dentales |
| 3. | Arrondissement des lèvres (y compris pour les consonnes suivies de voyelles labiales, par des phénomènes d'anticipation) |

Notons que d'autres particularités du français doivent faire l'objet de l'apprentissage phonétique, selon P. Delattre. Il s'agit de la :

- | | |
|----|---|
| 1. | Nasalité vocalique |
| 2. | Prévalence de l'intonation montante |
| 3. | Nature durative de l'accent primaire placé sur la syllabe finale |
| 4. | Loi de position pour la distribution des voyelles moyennes |
| 5. | Loi de trois consonnes pour déterminer l'emploi des voyelles ou semi-voyelles |

Un cours de phonétique devrait s'appuyer sur des matériaux méticuleusement construits, privilégiant la qualité à la quantité. Il devrait présenter les difficultés phonétiques graduellement, c'est-à-dire aborder d'abord la prosodie et les principes généraux d'articulation (antériorité, position de la langue, des lèvres, la loi de position), puis les sons individuels, mais toujours intégrés dans des unités de sens.

3.2 Difficulté d'apprentissage des voyelles

L'acquisition phonétique des voyelles semble présenter une difficulté majeure pour les apprenants, notamment dans l'apprentissage du FLE (Racine *et al.*, 2012). Plusieurs raisons peuvent expliquer cette difficulté.

1) **Nombre de qualités vocaliques** : Crothers (1978) qui travaille sur la base de 209 langues du Projet d'archivage phonologique de Stanford (UPSID) ne trouve que sept langues ayant plus de neuf qualités vocaliques de base. Si le français actuel possède dix différentes qualités de voyelles orales, c'est parce que les contrastes sont basés non seulement sur l'antériorité/postériorité mais aussi sur quatre degré d'aperture ainsi que sur la labialité, opposant les trois voyelles antérieures arrondies et étirées. Cette complexité du système phonologique met en évidence les difficultés phonéto-phonologiques que peut rencontrer un apprenant de FLE.

2) **Imprécision de la configuration articulaire** : Du fait de leur caractère continu (Stevens *et al.*, 1969), il est impossible de décrire avec précision aux apprenants la configuration des organes de la parole dans la production des voyelles, notamment dans le cas des voyelles non fermées (Kamiyama and Vaissière, 2009). Mis à part la voyelle [œ] sans constriction, toutes les voyelles correspondent à un rétrécissement du conduit vocal à un ou plusieurs endroits, mais l'aire de constriction au niveau de la langue est plus large que lorsqu'il s'agit de consonnes et elle est donc moins clairement définissable. Les lèvres peuvent prendre de différentes configurations : elles sont plus protruses dans la réalisation du [u] que pour celle du [y] (Riordan, 1977), et elles sont le plus protruses dans la réalisation de la voyelle nasale [ɔ̃].

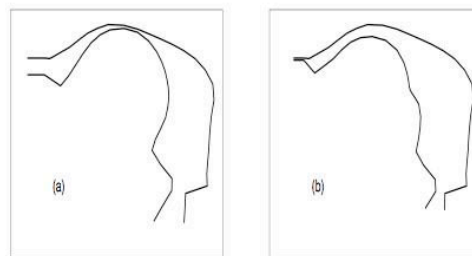
3) **Possibilités compensatoires** : Du fait des possibilités compensatoires entre les organes, un même timbre vocalique peut être généré par différentes configurations articulaires (Atal *et al.*, 1978; Maeda, 1989b). Par exemple, l'allongement de la cavité antérieure (qui résulte en un abaissement de la fréquence des formants principalement associés à cette cavité) peut être obtenu soit par un recul de la langue, soit par une protrusion, soit par un arrondissement des lèvres ou encore par les trois à la fois. Un recul de la langue provoquera un raccourcissement de la cavité postérieure (qui résulte en une élévation de la fréquence des formants principalement associés à la longueur de la cavité postérieure), mais ce raccourcissement pourra être compensé par un abaissement du larynx. Les gestes articulaires compensateurs sont flagrants chez un ventriloque ou un fumeur de pipe qui arrive à suppléer le mouvement de la mandibule par un mouvement plus ample de la langue (Vaissière, 2006).

Riordan (1977) a testé la possibilité d'atteindre la cible acoustique des voyelles labiales [u] et [y] sans arrondir les lèvres. Six locuteurs (quatre Français et deux Chinois Mandarin) ont enregistré les voyelles [i, y, u, ə, ɛ] en conditions normales et avec une bande élastique sur la bouche afin d'empêcher les lèvres de s'arrondir. Les valeurs basses de F1 et F2 pour [u] et de F3 pour [y],

nécessaires pour la perception du timbre souhaité, peuvent être assurées par l'action de la langue et du larynx dont les mouvements sont observés à l'aide d'un outil photoélectrique. Ainsi, en conditions normales, les formants moyens F1 et F2 de la voyelle [u] sont respectivement de 283 Hz et 733 Hz alors que quand les lèvres sont étirées par une bande élastique, les formants moyens sont de 300 Hz et 750 Hz. Dans le cas de [y], les formants moyens F1, F2 et F3 en conditions normales sont respectivement de 241 Hz, 1767 Hz et 2283 Hz alors qu'avec la bande élastique, ils sont de 250 Hz, 1791 Hz et 2283 Hz. Le résultat montre donc que la protrusion des lèvres n'est pas nécessaire pour la production de la cible acoustique [u] et [y].

De même, Bell-Berti *et al.* (1979) ont démontré la possibilité d'atteindre la même cible acoustique des voyelles antérieures de l'anglais [i, ɪ, e, ɛ] avec deux stratégies articulatoires distinctes, en manipulant soit la hauteur de la langue soit la tension de la langue.

Ouni and Laprie (2001) ont modélisé les voyelles arrondies [y, u] du français, en se basant sur le modèle articulatoire de S. Maeda. Ils ont démontré que le timbre de [y] (Figure 25) peut être généré soit par les lèvres plus ouvertes et une constriction plus postérieure (image a), soit par les lèvres plus fermées et une constriction plus antérieure (image b).



Two vocal tract configurations of the french vowel [y]. (a) the place of the constriction is in a rear position compared to (b). However, in (a) the lips are wide open and in (b) they are slightly open.

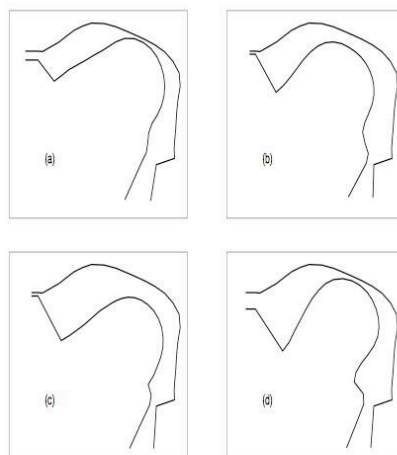
Figure 25 : Modélisation de [y] par Ouni and Laprie (2001)

Le même timbre [u] (Figure 26) est ensuite modélisé par des configurations du conduit vocal différentes. La constriction est alors soit vélaire avec les mâchoires plus écartées (image a) ou vélaire avec les mâchoires plus serrées (image d), soit palatal (image b), soit pharyngale (image c).

L'existence de stratégies articulatoires diverses dans la production des sons de la parole, notamment dans la production des voyelles, montre l'insuffisance des méthodes phonétiques qui reposent sur des descriptions purement articulatoires. Ainsi, selon Borrell (1993, p. 101), « *le but est donc de produire des sons à une certaine qualité acoustique sans se préoccuper de savoir quels sont les gestes articulatoires qui permettent de les réaliser.* »

4) Élasticité des voyelles : Les voyelles sont caractérisées par une grande élasticité (Duez, 2001) : le nombre de phonèmes vocaliques dans une langue peut varier, selon les régions, et dans le temps, alors que celui des consonnes est stable (Vaissière, 2006; Volin and Studenovsky, 2007). Par conséquent, selon les variétés du français (du Nord ou du Sud, par exemple) à laquelle l'apprenant est exposé durant son apprentissage, il peut rencontrer des qualités vocaliques différentes lors de la réalisation d'un même phonème. Certains contrastes entre voyelles peuvent être neutralisés dans une des variétés. Alors que les Français les plus âgés du Nord distinguent entre *pattes* /pat/ et *pâtes* /pat/, la jeune génération quelque soit son origine géographique ne fait plus cette distinction. De même alors que les mots *Baule* et *bol* sont prononcés de la même manière dans le Sud, les deux qualités [o] et [ɔ] correspondent à deux phonèmes /o/ et /ɔ/ dans une autre région (Vaissière, 2006).

3. Apprentissage phonétique



Different vocal tract configurations of the vowel [u]. (a) velar [u], (b) palatal [u] and (c) pharyngeal [u]. This classification is a function of the place of constriction. the configuration (d) is a variety of the configuration (a) where the jaw position is not the same but both produce the same vowel [u].

Figure 26 : Modélisation de [u] par Ouni and Laprie (2001)

5) **Caractère identique** : Les méthodes d'enseignement de LE se mettent d'accord que les sons de la LE *identiques* aux sons de la LM ne posent aucune difficulté phonétique car ils s'acquièrent par un simple transfert positif (Weinreich, 1953). Alors que plusieurs études interlangues reportent le caractère identique de certaines consonnes dans des langues différentes, notamment des nasales [m, n] (Wode, 1994), bon nombre d'auteurs généralement hésitent quant à l'emploi de ce terme dans la description des voyelles qui ne sont alors que très rarement décrites comme *identiques* (Bohn and Flege, 1990; Strange *et al.*, 2005).

6) **Conflit entre les tendances générales et les règles particulières** : Le français se caractérise par une prédominance de syllabes ouvertes, renforcée par les phénomènes de liaison et d'enchaînement vocalique (Delattre, 1953) où il existe une tendance pour les voyelles moyennes à utiliser le timbre fermé. Ainsi, il est de bon usage de distinguer entre 'chanterai' et 'chanterais', mais la tendance générale pointe vers l'utilisation du timbre [e] en syllabe finale des deux mots. Ainsi, l'apprenant se trouve confronté au conflit entre cette tendance générale et les règles particulières.

7) **Orthographe** : Dans le cas des méthodes qui utilisent des supports écrits, la difficulté à prononcer certaines voyelles du français est également due à l'orthographe qui serait moins divergente pour les consonnes (Léon, 1993). L'orthographe du français est selon Delattre (1960, p. 490) : « *irrationnelle parce qu'elle présente plus de fautes d'étymologie qu'on n'en peut compter, illogique parce qu'elle force à écrire tant de lettres qui ne se prononcent pas* » et elle est en fait « *un masque qui cache la réalité de la langue* ». Pour cette raison, Delattre (1944) sollicite les enseignants de LE à introduire la lecture et l'écriture d'une phrase une fois qu'elle est phonétiquement maîtrisée avec authenticité.

Les sept points développés ci-dessus expliquent pourquoi nous nous intéressons aux limites de l'acquisition par des non-natifs des voyelles plutôt que des consonnes.

Notons que nous ne traçons pas dans ce travail de la prosodie quoiqu'elle soit aussi importante à maîtriser, si l'on veut 'faire français', que les segments.

3.3 Etudes antérieures sur l'apprentissage phonétique des voyelles du français

Les études sur l'apprentissage phonétique du français sont nombreuses et nous en citons quelques-unes.

Enoch (1967) effectue une série de tests de discrimination par des apprenants israéliens parlant l'hébreux israélien (le nombre n'est pas précisé) de 16 voyelles du français [i, e, ε, a, α, ɔ, o, u, y, ø, œ, ə, ẽ, ã, õ, õ̃] et de 20 (semi)-consonnes [j, ɥ, w, p, t, k, b, d, g, f, s, ʃ, v, z, ʒ, l, ʁ, m, n, ɲ], insérées dans des mots. Il présente 3 mots à la suite (par exemple *vie, vue, vue*) et demande d'indiquer les mots prononcés de la même façon. Le résultat montre que les sons les plus facilement discriminés sont les voyelles [a, u, i] avec respectivement 0,7 %, 1,2 % et 1,5 % d'erreurs et les consonnes [m, ʃ, d, n] avec respectivement 2,3 %, 3,1 %, 3,6 % et 4 % d'erreurs. En revanche la voyelle [α] n'a jamais été "correctement" discriminée (100% d'erreurs) et les voyelles antérieures arrondies et nasales [y, ẽ, ã, œ, ø] ont été discriminées avec respectivement 26,8 %, 25,8 %, 23,2 %, 23,2 % et 22,7 % d'erreurs. Les consonnes les plus difficilement discriminées étaient en fait des semi-consonnes [ɥ, w] (avec respectivement 19,9 % et 18,7 % d'erreurs). Ainsi, les voyelles du français présentent un pourcentage d'erreurs de discrimination plus important que les consonnes.

Flege and Hillenbrand (1984) s'intéressent à l'identification par 7 Français natifs de deux voyelles du français [y, u], produites en contexte tVl par deux groupes de 7 apprenants américains, débutants et avancés dans 3 tâches : lecture de phrases, production spontanée de phrases qui commencent par « Tous les » ou « Tu les » et narration d'une histoire (au total 210 stimuli : 7 locuteurs dans chaque groupe * 2 voyelles * 3 tâches de production * 5 répétitions). Les résultats perceptifs montrent que la suite [tu] produite par les 7 apprenants débutants n'a été "correctement" identifiée par les 7 Français natifs que dans 45 % des réponses alors que [tu] produit par les 7 apprenants avancés a été "correctement" identifié dans 83 % des réponses ; il existe alors une différence significative entre les deux groupes d'apprenants selon leur niveau dans la production de [u] qui est considérée comme une voyelle *similaire*. En revanche, il n'existe pas de différence significative dans la production de la suite [ty] entre les apprenants débutants (identification correcte dans 70 %) et avancés (identification correcte dans 72 % des réponses). Les auteurs concluent que l'apprentissage d'une voyelle *similaire* [u] se fait progressivement alors que l'acquisition d'une voyelle *nouvelle* [y] est un processus relativement rapide.

Flege (1987d) s'intéresse à la production des voyelles [y] et [u] (mesure des formants F1-F3) et de l'occlusive [t] (mesure de VOT) produites dans des phrases qui commencent par « tous les » ou « tu les » par trois groupes d'Américains : 7 étudiants du français, 7 professeurs du français et 7 Américains habitant Paris. Les valeurs formantiques de F2 des voyelles françaises [y, u] et le VOT de l'occlusive française [t] produites par les Américains sont comparées avec les valeurs des sons produits par 7 Français natifs (« la référence »). Le résultat montre que la voyelle *nouvelle* [y] est produite de manière authentique (avec un F2 semblable à celui des natifs) par les Américains alors que le F2 de la voyelle *similaire* [u] produite par les trois groupes d'Américains n'atteint pas les mêmes valeurs que le F2 des Français natifs. De même, le VOT de la consonne *similaire* [t] produite par les trois groupes d'Américains est significativement différent de celui des natifs. L'auteur conclut que les voyelles *similaires*, comme [u], ne peuvent pas être réalisées d'une manière authentique malgré de nombreuses années d'expérience car leur acquisition est bloquée par le mécanisme de « classification par équivalence » alors que l'acquisition authentique des sons *nouveaux* est possible.

Gottfried and Beddor (1988) mènent une étude sur l'identification des voyelles moyennes /o, ɔ/ du français par 12 apprenants anglophones américains. Le test d'identification avec une échelle d'évaluation de 1 à 5 est construit à partir d'un continuum synthétisé, allant de [kot] à [kɔt]. Ainsi 10 stimuli sont créés en variant le F1 au début de la voyelle (par pas de 5 Hz, allant de 315 à 360 Hz), au milieu vocalique (par pas de 15 Hz, allant de 400 à 535 Hz) et à la fin vocalique (par pas de 5 Hz,

allant de 260 à 305 Hz) de même qu'en variant le F2 au début de la voyelle (par pas de 15 Hz, allant de 1270 à 1405 Hz), au milieu vocalique (par pas de 35 Hz, allant de 800 à 1115 Hz) et enfin en finale vocalique (par pas de 15 Hz, allant de 1340 à 1475 Hz). Chacun des 10 stimuli rencontre en plus 3 variations temporelles : les voyelles longues avec une durée de 220 ms, les voyelles moyennes de 180 ms et les brèves de 140 ms. Pour comparaison, les stimuli sont également identifiés par 12 Français natifs. Le résultat montre que la durée vocalique n'est pas perceptivement importante pour les Français natifs dans la distinction des voyelles [o, ɔ], alors qu'elle l'est pour les Américains. Ainsi, dans le cas d'ambiguïté spectrale, les voyelles longues sont identifiées par les Américains en tant que /o/ alors que les voyelles brèves sont identifiées en tant que /ɔ/.

Rochet (1995) étudie la production et la perception du [y] français par 10 apprenants anglophones du Canada et 10 apprenants lusophones du Brésil. La tâche de production consiste en répétition d'après un Français natif d'une liste de monosyllabes composées de voyelles [i, y, u, a] insérées dans des contextes consonantiques variés. Les enregistrements ont servi de stimuli pour un test d'identification par 3 Français natifs qui ont évalué les voyelles sur une échelle de 1 à 7. Le résultat de la tâche de perception montre que dans 52 % des réponses, les voyelles de mots reproduits par les apprenants lusophones ont été "correctement" identifiées par les 3 Français natifs en tant que /y/. Dans les 48 % de réponses restantes, les voyelles ont été majoritairement identifiées en tant que /i/ (dans 95 % des réponses), puis en tant que /u/ (dans 5 % des réponses). Les apprenants anglophones quant à eux ont reproduit des mots comportant des voyelles "correctement" identifiées en tant que /y/ dans 51 % des réponses des natifs. Dans les 49 % des réponses restantes, les stimuli ont été identifiés en tant que /u/ (dans 92 % des réponses) et en tant que /i/ (dans seulement 8 % des réponses). La tâche de perception consistait en identification par les apprenants d'un continuum vocalique allant de [i] à [u] (d'une longueur de 200 ms, avec le F1 constant de 250 Hz et le F2 variant par pas de 100 Hz, allant de 500 Hz à 2500 Hz) au travers de 2 catégories /i/ et /u/. Le résultat est en accord avec celui de la tâche de production : les stimuli avec un F2 entre 1300 Hz et 1900 Hz ont été identifiés en tant que /u/ par les apprenants anglophones et en tant que /i/ par les apprenants lusophones. Par cette étude, l'auteur montre l'importance de l'influence de la LM sur la perception et production de la LE.

Birdsong (2003) s'intéresse à la possibilité d'acquisition authentique par 22 apprenants anglophones tardifs des sons individuels du français (voyelles [i, e, o, u] et consonnes [p, t, k]). L'authenticité est alors estimée à partir de mesures de durée vocalique et du VOT, et au taux d'authenticité globale, noté sur une échelle de 1 à 5 par 3 juges français natifs. Du point de vue de production, les sons qualifiés d'authentiques produits par les apprenants sont ceux dont les valeurs de durée pour les voyelles et de VOT pour les consonnes se trouvent à moins d'un écart type des valeurs moyennes calculées à partir de productions de 17 natifs. Le résultat montre que plusieurs apprenants produisent les sons authentiquement ce qui suggère que la prononciation du français à la native n'est pas exclue dans le cas des apprenants anglophones tardifs.

Levy and Strange (2008) effectuent un test de discrimination catégorielle AXB par dix apprenants anglophones américains des contrastes vocaliques du français u/y, u/œ, i/y et y/œ, enregistrées par trois natives dans deux contextes consonantiques rabVp et radVt. Le contraste le plus difficile pour les Américains s'avère u/y (avec 30 % d'erreurs). Le contexte consonantique n'a pas d'effet significatif sur la capacité à discriminer la voyelle malgré une tendance vers une meilleure performance en contexte labial.

(Kamiyama and Vaissière, 2009) examinent la production des voyelles isolées [u], [y] et [ø] par 3 apprenants japonophones. L'étude acoustique dévoile une difficulté maximale dans la production du [u] qui est réalisé par les 3 japonophones avec un F2 supérieur à 1000 Hz et qui est perçu par les Français natifs en tant que [ø]. Le [ø] est de difficulté moyenne car il est produit par 2 apprenants avec une focalisation du deuxième et troisième formant et seulement un apprenant arrive à produire cette voyelle avec un F2 de 1500 Hz. Enfin le [y] est maîtrisé par 2 apprenants qui rapprochent les deuxième et troisième formants. Les valeurs formantiques moyennes des voyelles des natifs ont été calculées à partir de productions de quatre Français.

Expérimentateur(s)	Etude	Corpus	LM	Locuteurs	Auditeurs	Résultat
Enoch (1967)	Perceptive discrimination)	(de - 16 voyelles du français [i, e, ε, a, ɑ, o, u, y, ø, œ, ə, ɛ̃, ɑ̃, œ̃] insérées dans des mots - 20 (semi)-consonnes [j, ɥ, w, p, t, k, b, d, g, f, s, ʃ, v, z, ʒ, l, ʁ, m, n, ɲ]	Hébreux israélien	Non précisé	Apprenants israéliens (le nombre pas précisé)	- voyelles plus difficiles à discriminer que les consonnes - [ɑ] jamais “correctement” discriminée (100 % d’erreurs) - [y, ɛ̃, œ̃, œ, ø] discriminées avec difficulté
Flège and Hillenbrand (1984)	Perceptive (d’identification)	- [y, ɥ] du français en contexte tVl	anglais	2 groupes de 7 apprenants américains : débutants et avancés	7 Français natifs	- [tu] produit par les apprenants avancés mieux identifié par les natifs que [tu] produit par les débutants - [ty] produit par les avancés aussi bien identifié que [ty] produit par les débutants ⇒ apprentissage de la voyelle <i>similaire</i> [u] est progressif alors que celui de [y] se fait rapidement
Flège (1987d)	De production (mesure de F1-F3 et de VOT)	- [y, ɥ] et [t] du français en contexte tVl	anglais	- 3 groupes d’Américains : 7 étudiants du français, 7 professeurs du français et 7 Américains habitant Paris - 7 Français natifs		- voyelle <i>nouvelle</i> [y] authentique pour les 3 groupes (F2 semblable à celui des natifs) - voyelle <i>similaire</i> [u] pas authentique (F2 diffère de celui des Français natifs) - consonne <i>similaire</i> [t] pas authentique (VOT diffère de celui des natifs) ⇒ possibilité d’acquisition authentique des sons <i>nouveaux</i> mais pas des sons <i>similaires</i>
Gottfried and Beddor (1988)	Perceptive (d’identification)	- [o, ɔ] français : continuum synthétisé, allant de [kot] à [koʔ] avec variation temporelle	anglais	Continuum synthétisé	- 12 apprenants américains - 12 Français natifs	- durée vocalique joue un rôle dans la distinction des voyelles /o, ɔ/ par les Américains mais pas par les Français natifs - en cas d’ambiguïté spectrale, les sons vocaliques de longue durée sont identifiées par les Américains comme /o/ et les sons vocaliques brefs comme /ɔ/
Rochet (1995)	Perceptive (d’identification)	- [i, y, u, a] insérées dans des contextes consonantiques variés (monosyllabes) - continuum vocalique allant de [i] à [u] avec durée constante	- anglais - portugais	- 10 apprenants anglophones du Canada - 10 apprenants lusophones du Brésil Continuum synthétisé	3 Français natifs - 10 anglophones du Canada - 10 lusophones du Brésil	- les Français natifs catégorisent le [y] des lusophones en tant que /y/, puis /i/ - les Français natifs catégorisent le [y] des anglophones en tant que /y/, puis /u/

3. Apprentissage phonétique

Birdsong (2003)	De production (mesure de durée vocale et du VOT)	- voyelles [i, e, o, u] en syllabe finale ouverte dans des mots mono/ bisyllabiques	anglais	- 22 apprenants anglophones tardifs - 17 Français natifs (dont 7 femmes)		- plusieurs apprenants produisent les sons authentiquement (avec les valeurs mesurées à moins d'un écart type de celles des natifs)
	Perceptive (mesure de l'authenticité globale, échelle 1-5)	- consonnes [p, t, k] en syllabe initiale dans des mots mono/bi/trisyllabiques			3 Français natifs	- les productions de plusieurs apprenants jugées comme authentiques => possibilité de parler de manière native même lors d'un apprentissage tardif
Levy and Strange (2008)	Perceptive (de discrimination catégorielle AXB)	- contrastes vocaliques u/y, u/æ, i/y et y/æ, en contextes rabVp et radVt	anglais	3 Françaises natives	10 apprenants anglophones américains	- le contraste le plus difficile u/y (30 % d'erreurs) - pas d'effet significatif du contexte sur la discrimination des voyelles - tendance à une meilleure discrimination en contexte labial
Kamiyama and Vaissière (2009)	De production (mesure des formants F1-F3)	- voyelles isolées [u], [y] et [ø]	japonais	- 3 apprenants japonophones - 4 Français natifs (2 femmes, 2 hommes)		- difficulté (du + au -) : [u] -> [ø] -> [y]
	Perceptive (de discrimination AXB)	- contrastes vocaliques u/y, y/ø, u/ø, i/e, u/o, e/a		3 Français natifs (2 femmes, un homme)	14 apprenants japonophones	- difficulté (du + au -) : u/ø -> y/ø -> u/y => [y] français, phonémiquement et phonétiquement <i>nouveau</i> et [ø] phonémiquement <i>nouveau</i> plus facile que [u], phonémiquement <i>similaire</i> (au /u/ japonais, réalisé [ɯ]) et phonétiquement <i>nouveau</i>
Darcy <i>et al.</i> (2012)	- tâche de représentation lexicale	- 40 mots et 40 non-mots avec les contrastes vocaliques y/u et æ/ɔ formant des paires minimales et 120 mots distracteurs	anglais	2 Françaises natives	2 groupes d'apprenants américains : 20 du niveau avancé et 38 du niveau intermédiaire	- les apprenants avancés établissent des catégories phonologiques pour les voyelles des trois contrastes mais ils font des erreurs au niveau phonétique - les apprenants intermédiaires confondent les voyelles y/u au niveau lexical mais maîtrisent ce contraste dans la tâche de discrimination
	- discrimination perceptive AXB	- contrastes vocaliques y/u et æ/ɔ dans 48 non-mots CVC où C est une consonne labiale ou coronale				

Tableau 20 : Etudes antérieures sur l'apprentissage phonétique du français

3. Apprentissage phonétique

La perception de 14 apprenants japonophones est également examinée au travers d'un test de discrimination AXB composé de 72 triplets (6 paires u/y, y/ø, u/ø, i/e, u/o, ε/a * 4 combinaisons et ordre de présentation * 3 locuteurs). Le résultat du test montre que la distinction entre u/y est la plus facile (discriminés « correctement » dans 95 % des réponses) alors que celle entre u/ø s'avère la plus difficile (discriminés « correctement » dans 85 % des réponses). La distinction entre y/ø est de difficulté moyenne (discriminés « correctement » dans 91 % des réponses). Les auteurs concluent que la voyelle française [u] qui est phonémiquement *similaire* mais phonétiquement différente du [u] japonais (noté par le phonème /u/) est plus difficile à acquérir par les japonophones que la voyelle [y] (qui est phonémiquement et phonétiquement *nouvelle*) et que la voyelle [ø] (qui est phonémiquement *nouvelle* mais phonétiquement *similaire*). La Figure 27 résume leur conclusion.

	Difficulty in production →			
	/a/	/ø/	/y/	/u/
Phonemically similar	+	-	-	+
Phonetically similar	+	+	-	-

Figure 27 : Apprenants japonophones : niveau de difficulté de l'apprentissage phonétique des voyelles françaises [a, ø, y, u] selon leur similarité phonémique/ phonétique avec les voyelles du japonais selon Kamiyama and Vaissière (2009)

Darcy *et al.* (2012) étudient l'acquisition des contrastes vocaliques y/u et œ/ɔ par deux groupes d'apprenants anglophones américains, 20 du niveau avancé et 38 du niveau intermédiaire, dans une tâche de représentation lexicale (40 mots et 40 non-mots formant des paires minimales et 120 mots distracteurs) et de discrimination perceptive AXB (48 non-mots CVC où C est une consonne labiale ou coronale, les stimuli étant séparés avec un ISI⁹ de 500 ms). Les (non)-mots avaient été enregistrés par deux Françaises natives. La tâche de représentation lexicale reflète l'acquisition de catégories phonologiques alors que le test de discrimination reflète l'acquisition de catégories phonétiques. Les résultats montrent que les apprenants avancés établissent des catégories phonologiques pour les voyelles des trois contrastes mais qu'ils font des erreurs au niveau phonétique. Les apprenants intermédiaires, eux, confondent les voyelles y/u au niveau lexical mais maîtrisent ce contraste dans la tâche de discrimination. Les auteurs mettent en évidence une dissociation entre le niveau phonologique (lexical) et le niveau phonétique et montrent ainsi que l'apprentissage des deux types de catégories peut se réaliser indépendamment.

Le Tableau 20 (page 77) récapitule les différentes études effectuées en précisant l'expérimentateur, le type d'étude, le corpus, la langue maternelle des apprenants, les locuteurs, les auditeurs et le résultat.

3.4 Théories de l'apprentissage phonétique

Les théories traditionnelles de l'apprentissage phonétique, telles que la théorie de l'interférence et l'analyse contrastive, s'appuient sur la notion de phonème et de traits phonologiques pour expliquer les difficultés dans l'apprentissage d'un nouveau système sonore. En revanche, les théories plus récentes opèrent avec des unités concrètes (les sons) qui possèdent une caractéristique phonétique complexe qu'il faut prendre en compte (McAllister 1998, McAllister 2000). En sachant que chaque phonème a une multitude de réalisations possibles, les difficultés prédites par une analyse acoustico-perceptive sont plus détaillées que celles prédites par l'analyse contrastive. Parmi les modèles récents

⁹ ISI : Inter-stimulus interval

sont couramment cités le Speech Learning Model (SLM) de Flege (1995), le Perceptual Assimilation Model (PAM) de Best (1995) et le Native Language Magnet (NLM) de Kuhl and Iverson (1995).

3.4.1 Théorie de l'interférence et l'analyse contrastive

La théorie de l'interférence et l'analyse contrastive sont étroitement liées et émergent dans les années 1950-1960. Weinreich (1953) mène un important travail sur l'interférence : il la définit comme une déviation par rapport à la norme résultante du contact des langues. Il détermine trois types d'interférence : phonique, phonotactique et suprasegmentale et prédit les interférences en se basant sur les traits phonologiques.

De son côté, l'analyse contrastive permet de prédire les difficultés des apprenants en comparant le système phonologique de la langue maternelle et de la langue cible. Elle met en regard les inventaires phonologiques des deux langues ; les phonèmes qui n'existent pas dans la LM seraient plus difficiles à acquérir que les phonèmes existants et donc connus (Lado, 1957).

L'analyse contrastive pose de nombreux problèmes quant à la description de l'apprentissage phonétique d'une LE. Elle prédit soit un succès, soit un échec sans envisager l'existence d'étapes intermédiaires dans l'apprentissage (Flege, 1988b). En se basant sur les inventaires phonologiques, elle prédit des erreurs qui n'apparaissent pas systématiquement ou au contraire ne prédit pas certaines erreurs. En effet, toutes les erreurs ne peuvent pas être expliquées par les différences qui existent entre les systèmes phonologiques de la LM et de la LE. D'autres facteurs rentrent en jeu, notamment le type des sons présentés et la saillance perceptive des indices acoustiques (Burnham, 1986; Flege, 1989; Polka, 1991) (voir le chapitre 5 pour plus de précisions sur les facteurs de l'accent étranger). L'analyse contrastive, tout comme d'autres théories qui permettent d'identifier la source de l'erreur, reste néanmoins utile quand elle est réalisée a posteriori pour expliquer les difficultés (Major, 1987).

3.4.2 Speech Learning Model – SLM (Flege, 1995)

Le modèle de l'apprentissage phonétique SLM se base sur un ensemble d'hypothèses et de postulats, synthétisés par Flege (1995). Il a été élaboré à partir des résultats d'une multitude d'études de perception et production d'apprenants (de l'anglais ou d'autres langues), ayant des LM différentes (français, espagnol, chinois mandarin, etc.).

Selon l'auteur, au cours de l'acquisition de la LM, l'enfant crée des représentations mentales pour les catégories phonétiques, puis phonologiques de la langue ambiante. Les images mentales sont stockées dans la mémoire à long-terme et elles se reflètent en production. Ainsi, pour produire authentiquement un son de la LE, il faut d'abord posséder une représentation perceptive fidèle de ce son (Flege, 1997; Flege and Eefting, 1987). Notons que malgré ce postulat, l'auteur n'explique pas toutes les erreurs en production par la perception des apprenants (Harmegnies *et al.*, 2005).

Le processus de création ou de modification de catégories phonétiques ne s'arrête pas à l'âge adulte et les représentations peuvent ainsi évoluer au cours de la vie (Flege, 1981; Flege, 1987a). Pour qu'une nouvelle catégorie soit établie, il faut percevoir au moins quelques différences acoustiques entre le son de la LE et le son le plus proche de la LM (Flege and Eefting, 1987). Plus le son étranger est perceptivement proche du son connu, plus il est difficile de remarquer la différence. Ainsi, les sons perceptivement *similaires* présentent une difficulté maximale qui peut devenir insurmontable si l'apprentissage débute après l'âge de 5 à 7 ans (Flege, 1988a; Flege, 1991). En effet, dans la perception et production de ces sons (appelés diaphones), les apprenants « tardifs » activent une seule catégorie phonétique ce qui se manifeste par la présence de « l'accent étranger » (Flege, 1991). L'apprentissage est bloquée par le mécanisme de *classification par équivalence* (Flege, 1987a; Flege,

1987d; Flege, 1991; Flege, 1997) que Flege and Eefting (1987) décrivent également comme « old wine in new bottle ».

En revanche, l'apprentissage des sons *nouveaux* peut poser problème au début de l'apprentissage, mais au fur et à mesure de l'expérience, il est possible d'atteindre l'authenticité. Ainsi la différence essentielle entre l'apprentissage des sons *nouveaux* et celui des sons *similaires* est la suivante : les derniers semblent être plus faciles à acquérir, car on a l'impression de les connaître à cause de leur similarité phonétique avec les sons de la LM. Les enseignants y prêtent moins attention et les apprenants ignorent souvent les subtilités phonétiques qui les différencient des sons connus. A cause de cette apparente facilité, il y a peu de chance que les apprenants produisent et perçoivent les sons *similaires* comme les natifs (Flege and Eefting, 1987; Flege and Hillenbrand, 1984). En ce qui concerne l'acquisition des sons *nouveaux*, la situation est inverse : ils ne renvoient à aucun son de la LM et les apprenants ne possèdent pas de référence au début de l'apprentissage. Les *nouveaux* sons peuvent alors paraître difficiles car ils demandent des configurations articulatoires nouvelles et une cible acoustico-perceptive étrangère, voir étrange. Pour cette raison, tant qu'un prototype phonétique pour ces sons n'est pas créé, leur réalisation peut être très variable (Flege, 1991). Mais au fil de l'apprentissage, il est possible d'établir une nouvelle catégorie phonétique pour un son *nouveau* et même d'atteindre une performance native.

Les sons *identiques*, eux, ne posent a priori pas de problèmes pour l'apprentissage phonétique (Bohn and Flege, 1990; Flege, 1988b) car ils sont acquis par « transfert positif » des sons de la LM en LE.

Flege suggère également que l'apprentissage phonétique se déroule par étapes (Flege, 1980; Flege, 1987b) et l'apprenant se rapproche de la norme phonétique petit à petit en produisant des sons approximatifs. Il s'agit d'un processus lent où chaque étape correspond à une période de productions relativement stables, créant ainsi une interlangue (Flege, 1988b). Le progrès est difficile à observer tant qu'un son n'est pas produit authentiquement, d'autant que les stratégies d'apprentissage peuvent dépendre des individus. En effet, certains apprenants produisent des sons intermédiaires en utilisant des indices acoustiques différents des natifs (Bohn and Flege, 1990). D'autres peuvent créer un contraste en exagérant l'un des traits phonétiques, le plus souvent l'indice temporel, qu'ils jugent perceptivement saillant (Bohn and Flege, 1990; Flege, 1989). L'apprentissage phonétique peut aussi varier en fonction du son au sein de la même catégorie phonologique car certains allophones peuvent être acquis plus rapidement que d'autres (Flege, 1987a),(Levy, 2009).

Enfin, une personne bilingue utilise un espace phonétique commun pour les deux langues qu'elle parle. Afin de garder le contraste entre les sons, les représentations des catégories phonétiques de la LM et LE sont parfois modifiées (Flege, 1981). Ainsi, les bilingues peuvent avoir des images mentales différentes de celles des monolingues.

Pour résumer, le SLM est un modèle acoustique de perception et de production, particulièrement adapté aux apprenants avancés, ayant débuté l'apprentissage de la L2 après l'âge de 5-7 ans.

3.4.3 Perceptual Assimilation Model – PAM (Best, 1995)

Le Modèle PAM s'appuie sur une théorie écologique de perception qui revendique la possibilité de percevoir les sons de la parole directement, sans l'obligation de passer par des représentations mentales de ces sons. L'homme observe les gestes articulatoires de son interlocuteur, il voit directement le lieu d'articulation des sons produits et à partir de ces gestes, il lui est possible d'extraire immédiatement l'information linguistique (Best *et al.*, 2001). L'accès à l'information linguistique véhiculée par les gestes des organes de la parole est le résultat d'un long apprentissage où les enfants établissent progressivement un lien entre les constellations articulatoires spécifiques à leur

LM et le message linguistique qu'elles portent. Chaque langue posséderait ainsi des invariants linguistiques qui lui sont propres. Selon PAM, l'expérience linguistique avec la LM modifie rapidement la perception qui devient phonémique, c'est-à-dire qu'elle se réalise au travers des catégories phonologiques établies pour la LM. Ainsi suite à la réorganisation perceptuelle, les enfants entre 10 et 12 mois perçoivent certains contrastes non-natifs avec difficulté, par exemple les éjectifs, alors que ceux âgés de 6-8 mois ne présentent aucune difficulté (Best *et al.*, 1995; Best *et al.*, 1988).

Dans l'apprentissage d'une LE, il faut pouvoir créer une association entre les configurations articulatoires et le sens linguistique. Les configurations articulatoires qui ont une fonction phonologique dans la LM de l'apprenant sont naturellement plus facilement reconnues que les configurations nouvelles. Les sons de la LE sont souvent interprétés et perçus par les constellations gestuelles propres à la LM de l'apprenant qui sont les plus proches.

Malgré l'existence de configurations propres à une langue donnée, il existe également des constellations gestuelles communes. Selon Best, de nombreuses constellations sont partagées par les différentes langues du monde. Ceci est dû aux limites de l'appareil phonatoire de l'homme. Le modèle PAM se distingue de la théorie motrice dans le sens où elle ne stipule pas un module neuronal inné qui permet d'interpréter les configurations des articulateurs. Selon PAM, les auditeurs ne passent pas par les représentations motrices des gestes pour pouvoir les interpréter. Ainsi les sons (la phonétique) et le sens qu'ils véhiculent (la phonologie) émergent au même moment, à la sortie des organes de la parole.

Ce modèle fournit des prédictions sur l'assimilation perceptive des contrastes non-natifs. La perception des sons de la LE qui sont en contraste dépend de leur similarité phonétique et phonologique avec les sons de la LM (Best *et al.*, 2001). Selon PAM, la similarité phonologique fait qu'on utilise l'une ou l'autre catégorie native alors que la similarité phonétique provoque notre jugement de justesse sur la réalisation de cette catégorie. La similarité est définie dans PAM par des caractéristiques articulatoires. Ainsi, dès que possible, l'auditeur essaye d'assimiler le son non-natif au phonème natif qui est le plus proche au niveau articulatoire. En fonction du type de contraste, PAM prédit les assimilations possibles suivantes :

- les deux sons de la LE sont assimilés à une seule catégorie native ; les deux sons présentent le même degré de justesse que la catégorie native ; ils sont soit de bons soit de mauvais exemples. La discrimination attendue est mauvaise ;
- les deux sons de la LE sont assimilés à une seule catégorie native ; l'un est alors un bon représentant et l'autre un mauvais représentant de la catégorie native. La discrimination attendue est moyenne ;
- les deux sons de la LE sont assimilés à des catégories natives différentes. On attend une très bonne, voir excellente discrimination ;
- un son de la paire renvoie à une catégorie native, l'autre son ne peut pas être catégorisé. La discrimination devrait alors être très bonne ;
- les deux sons ne renvoient à aucune catégorie native. La discrimination est soit mauvaise (si les deux sons renvoient au même ensemble de sons maternels), soit moyenne ;
- les sons du contraste sont tellement différents au niveau articulatoire des phonèmes natifs que l'apprenant ne les identifie pas comme sons de la parole. Les catégories ne sont pas assimilables. La discrimination attendue est bonne, voir excellente car les sons échappent aux effets de la LM.

Le type d'assimilation le plus fréquent est celui qui lie les deux catégories non-natives à une seule catégorie native (Best *et al.*, 1995). Le type d'assimilation le plus rare se produit pour les sons que l'auditeur n'identifie pas comme sons de la parole (Best and Tyler, 2007). C'est par exemple le cas des Anglophones qui perçoivent les clics du zoulou (Best *et al.*, 2001) : même s'ils n'ont jamais entendu ces sons en anglais, ils peuvent les discriminer sans difficulté. En effet, en l'absence d'assimilation avec les sons de la parole, ils accèdent à la perception phonétique (Best *et al.*, 1988) et ils peuvent se concentrer par conséquent sur les détails articulatoires.

PAM prédit la hiérarchie de difficultés de perception des contrastes non-natifs (Best *et al.*, 2001). Les sons les plus difficiles à discriminer sont ceux qui correspondent à la même catégorie native, sans différence de jugement. Ensuite, les sons qui renvoient à la même catégorie native mais avec des degrés de justesse différents sont plus faciles. Enfin, les sons qui activent deux catégories natives distinctes sont les plus faciles à discriminer.

PAM se distingue de SLM et NLM par la prédiction de la facilité à discriminer un contraste entre deux sons de la LE qui correspondent à une seule catégorie native mais provoquent des jugements de justesse différents. Dans ce type d'assimilation, une paire de sons *similaires* peut être relativement facile à discriminer.

Best *et al.* (2001) ne soutiennent pas l'hypothèse de la période critique. Malgré une diminution de capacité à discriminer certains contrastes non-natifs avec l'âge, il est possible d'améliorer la performance avec des entraînements dans des laboratoires phonétiques ou avec une expérience naturelle.

Le modèle dans sa version d'origine explique la perception des contrastes non-natifs par des auditeurs naïfs qui n'ont aucune expérience avec la LE. Il ne concerne pas directement les apprenants de la LE et ne permet pas de prédire les différentes étapes d'apprentissage. C'est pour cette raison que Best and Tyler (2007) font évoluer le modèle en créant PAM-L2.

3.4.3.1 PAM-L2 (Best and Tyler, 2007)

Ce modèle est dédié spécifiquement à l'apprentissage d'une LE et prédit le succès pour un apprentissage perceptif de la LE chez les apprenants. PAM-L2 s'adresse notamment aux apprenants ayant appris la LE d'une manière naturelle, c'est-à-dire dans une situation d'immersion. Les apprenants sont ici définis comme des sujets qui apprennent activement la LE, dans le but de pouvoir communiquer. Leur apprentissage est influencé par de nombreux facteurs qui interagissent ; il s'agit de l'âge au début de l'apprentissage, qualité et quantité d'input, durée de résidence dans le pays où l'on parle la LE et utilisation relative de la LE par rapport à la LM.

PAM-L2 prévoit les évolutions suivantes :

- quand les deux sons de la L2 renvoient à un son de la L1 mais avec des degrés de justesse différents, une nouvelle catégorie peut être créée pour le mauvais représentant mais pas pour le bon. La nouvelle catégorie sera d'abord d'ordre phonétique, puis éventuellement d'ordre phonologique ;
- quand les deux sons de la L2 correspondent à un seul son de la L1 et qu'ils présentent le même degré de justesse (ils sont aussi bons ou aussi mauvais l'un que l'autre), ces sons seront difficiles à discriminer au début et ils seront alors perçus comme homophones. La discrimination pourrait s'améliorer pour les sons qui sont des mauvais représentants de la catégorie native mais ce progrès n'est pas réellement attendu par PAM. La chance de progresser augmente avec la fréquence des mots : si les mots contenant ces sons sont très fréquents dans le vocabulaire et si les sons permettent de créer de nombreuses paires minimales dans la L2, il y a alors plus de chance de réussite pour un apprentissage perceptif ;
- quand seulement l'un des deux sons est assimilé perceptivement à une catégorie native et l'autre son n'est pas catégorisé, la discrimination est bonne. Le son assimilé peut être perçu comme un bon ou un mauvais exemple de la catégorie native. Si c'est un mauvais représentant, une nouvelle catégorie pourra être créée. S'il s'agit d'un bon exemple, il y a peu de chance de créer une nouvelle catégorie pour ce son ;
- quand les deux sons ne sont assimilés à aucune catégorie native, ils renvoient à des catégories natives multiples où aucune ne prédomine. Dans les cas où les deux sons renvoient à des ensembles de sons maternels différents, les catégories phonologiques de L2 peuvent être faciles

à acquérir. Si les deux sons de L2 renvoient en revanche au même ensemble de sons, une seule catégorie phonologique sera créée pour ces deux sons ;

- quand les deux sons ne sont pas perçus comme des sons de la parole, leur discrimination est bonne dès le départ même chez des auditeurs naïfs.

Pour résumer, pour les sons qui sont assimilés à une catégorie native avec un grand degré de justesse, aucun progrès significatif n'est prévu. Ainsi, les voyelles qui sont de bons représentants de sons maternels ne subiront pas d'importants changements lors de l'apprentissage. En revanche, PAM-L2 prévoit un apprentissage pour les voyelles qui sont perçues comme différentes des voyelles natives. Le succès de la création de nouvelles catégories phonétique et phonologique pour les deux sons en contraste dépendra de la similarité perçue entre elles. Si les deux sons renvoient au même ensemble de sons *similaires*, la différence perçue entre les deux sons est alors moindre. Si en revanche les deux sons renvoient à des ensembles de sons différents, la différence perçue entre les deux sons est plus grande et les deux sons sont alors plus faciles à discriminer.

PAM est alors un modèle articulatoire de perception, qui prédit l'évolution de l'apprentissage depuis le niveau débutant (PAM) au niveau plus avancé (PAM-L2).

3.4.4 Native Language Magnet – NLM (Kuhl and Iverson, 1995)

La théorie NLM, basée sur de nombreuses études, réalisées avec des adultes, des enfants et des animaux, décrit le développement de la perception depuis l'enfance jusqu'à l'âge adulte.

Les expériences de Kuhl démontrent l'influence de la langue ambiante sur la perception. La langue parlée autour de nous agit sur notre perception même avant la naissance (Kuhl, 2000). En effet, à sa naissance, l'enfant marque une préférence pour la voix de sa mère et pour les comptines qui lui ont été racontées au cours des dix dernières semaines dans l'utérus.

NLM décrit trois états dans le développement de la perception. Au tout début de la vie, les nourrissons présentent une capacité phonétique universelle. Ils sont capables, grâce au mécanisme d'audition général (Kuhl 1991), de discriminer les différents sons des langues du monde et pas seulement ceux de leur LM. L'espace vocalique est alors divisé par des frontières naturelles. Au fur et à mesure de l'expérience avec la LM, les frontières des catégories phonétiques sont redistribuées (état 2). Les enfants, à partir de l'âge de 6 mois, créent des représentations mentales pour les sons de leur LM ce qui déforme leur perception pour les sons non-natifs. Cette capacité est propre à l'homme. Les représentations stockées en mémoire sont les prototypes ou bien les meilleurs représentants des sons. Ils sont le résultat abstrait d'un calcul statistique sur tous les sons entrants. Les sons qui se trouvent en proximité du prototype, sont attirés par ce dernier et les différences perceptives à l'intérieur de chaque catégorie sont minimales (Iverson *et al.*, 2001; Kuhl and Iverson, 1995). Les stimuli qui se situent aux frontières entre les catégories sont perçus comme différents du prototype. Les prototypes se comportent en effet comme des aimants, attirant les unités qui se trouvent dans leur champ magnétique. La perception finit par être complètement déformée (état 3) par les prototypes établis pour les sons de la LM.

Les catégories de la LM exercent le même pouvoir d'aimant sur les sons d'une LE. Les sons *similaires* sont attirés et interprétés par le prototype existant pour la LM alors que les sons *nouveaux* ne sont généralement pas associés aux catégories de la LM (Harmegnies *et al.*, 2005).

Pour expliquer les changements de perception, Kuhl observe l'activité cérébrale (Kuhl *et al.*, 2008). L'exposition à la langue ambiante crée des changements dans les tissus des neurones. Au fur et à mesure de l'exposition de l'enfant à la LM, les représentations neuronales pour les catégories natives sont formées. Elles permettent de porter l'attention sur les patterns de la L1 mais diminuent la sensibilité envers les schémas phonétiques étrangers. Plus les représentations neuronales sont stables,

3. Apprentissage phonétique

plus il est difficile de créer de nouvelles catégories pour les sons étrangers. Même un séjour dans le pays où l'on parle la LE ne permet pas de modifier les représentations neuronales établies.

Pour résumer, il s'agit ainsi d'un modèle de perception qui prédit la difficulté phonétique des apprenants au tout début de l'apprentissage (niveau débutant) où ils peuvent encore être considérés comme auditeurs naïfs, n'ayant pas de connaissance de la L2.

Les trois modèles récents développés ci-dessus traitent alors essentiellement des sons (voyelles et consonnes) en évitant les phénomènes prosodiques. Ils ne soutiennent pas l'hypothèse de la période critique mais montrent l'influence de la langue maternelle sur la langue étrangère. Alors que PAM et NLM s'intéressent exclusivement à la perception des sons, SLM prédit également les difficultés que peuvent rencontrer les apprenants en production. En effet, c'est le seul des trois modèles qui considère que toutes les erreurs en production ne peuvent pas être clarifiées en termes de la perception erronée de l'apprenant. Le Tableau 21 récapitule les postulats de base de ces trois modèles.

Nom (auteurs)	Type	Apprenants visés	Difficulté prévue	
			+difficile	+facile
SLM (Flege, 1995)	Modèle acoustique : signal acoustique -> représentations mentales -> sens	Avancés	Sons <i>similaires</i> : « classification par équivalence », les différences acoustiques non perçues	Sons <i>identiques</i> Sons <i>nouveaux</i> : la différence est perçue rapidement (catégorie perceptive créée), problème plutôt au niveau production mais surmontable >> acquisition authentique possible
PAM (Best, 1995), (Best and Tyler, 2007)	Modèle articulatoire direct : gestes articulatoires -> sens	Locuteurs naïfs = apprenants débutants (Best, 1995) ; PAM- L2 : apprenants avancés (Best and Tyler, 2007)	2 catégories assimilées à une catégorie native (mais la difficulté à percevoir le contraste dépend du degré de justesse)	- 2 catégories assimilées à 2 catégories natives - 1 son assimilé, 1 son non catégorisé - les 2 sons non assimilables (clics pour les anglophones) - possibilité de créer une nouvelle catégorie pour les mauvais représentants, pas pour les bons
NLM (Kuhl and Iverson, 1995), (Kuhl <i>et al.</i> , 2008)	Modèle acoustique : signal acoustique -> représentations mentales (prototypes) -> sens	Locuteurs naïfs = apprenants débutants	Sons <i>similaires</i> : à proximité du prototype établi pour une catégorie native (effet d'aimant)	Sons <i>nouveaux</i> : échappent à l'effet d'aimant car loin du prototype

Tableau 21 : Modèles récents de l'apprentissage phonétique

4. Enseignement de la phonétique

UNE ILLUSION TROP REPENDUE est qu'il suffit de savoir parler et écrire une langue pour être capable de l'enseigner.

(Delattre, 1960, p. 483)

Résumé : Le **quatrième chapitre** est consacré à l'enseignement de la phonétique qui demande à l'enseignant une formation et des connaissances approfondies. Ce dernier s'appuie souvent sur des outils traditionnels tels l'Alphabet Phonétique International (API) et le triangle vocalique ou encore sur des outils plus récents comme les modèles articulatoires de production de la parole ou les spectrogrammes en temps réel. Nous introduisons la notion de « norme phonétique » dont la définition est nécessaire dans les études instrumentales (Flege, 1987b).

Les méthodes conventionnelles d'enseignement des LE, telles que la méthode traditionnelle, directe, audio-orale ou communicative sont passées en revue et nous mettons l'accent sur la place qu'elles accordent à la phonétique. Les méthodes visant la correction phonétique sont notamment la méthode verbo-tonale, la méthode articulatoire et la méthode des oppositions phonologiques. La première prend en compte la coarticulation et la compensation articulatoire en insistant sur les caractéristiques acoustico-perceptives des sons. La seconde préconise la description articulatoire de chaque son alors que la troisième donne de l'importance aux exercices structuraux basés sur des unités sémantiques. Notons que les deux dernières sont complémentaires.

Enfin, nous abordons l'évaluation du niveau phonétique des apprenants dont la production s'améliore progressivement mais avant d'atteindre la cible telle qu'elle est réalisée par un natif, il est difficile d'observer le progrès. Les critères phonétiques retenus par le Cadre Européen Commun de Référence pour les langues (Conseil de l'Europe, 2001) pour définir les six niveaux de compétences des apprenants reflètent la difficulté à évaluer avec exactitude le niveau phonétique.

L'enseignement de la phonétique d'une LE est jugé difficile (Blake, 1972; Wachs, 2011). Il s'agit d'une matière sensible et parfois taboue (Billières and Borrell, 1990) au point que les enseignants évitent souvent les exercices de phonétique, du moins avec les apprenants débutants (Fernandez, 2012) comme en témoigne le terme de « correction phonétique » (Harmegnies *et al.*, 2005). Il est pourtant admis que la fossilisation des erreurs est dangereuse (Galazzi-Matasci, 2011; Saunders, 2009) car il est difficile de modifier certaines habitudes articulatoires établies (Flege, 1980).

Si l'enseignement phonétique est estimé difficile, c'est parce qu'il demande une longue préparation et de nombreuses connaissances de la part de l'enseignant. Selon Delattre (1944, p. 114), ce dernier devrait :

pouvoir anticiper les difficultés d'articulation et de diction des élèves. Pour cela il faut savoir décrire la langue de l'élève et la langue qu'on lui enseigne tant du point de vue physiologique (positions, mouvements des organes de la parole) que du point de vue acoustique (timbre des sons, syllabation, rythme, intonation, place et nature de l'accent, etc.). Il faut aussi connaître l'essentiel de l'histoire des lois et tendances qui gouvernent la prononciation moderne de la langue à apprendre.

L'enseignant doit en outre permettre à l'apprenant de construire son propre système phonétique à partir de la norme (Guimbretière, 1994) ce qui s'avère délicat en milieu exolingue où l'enseignant n'est pas toujours un locuteur natif de la langue et présente lui-même des déviations par rapport à la norme (Billières and Borrell, 1990).

Ainsi l'absence des connaissances en phonétique et d'un modèle chez les enseignants d'une LE, fréquemment amateurs dans ce domaine, entraîne des résultats souvent médiocres en diction et prononciation de la L2 par leurs élèves.

4.1 Quelques outils de l'enseignement phonétique

L'enseignant de la phonétique d'une LE s'appuie souvent sur l'Alphabet Phonétique International (API) et le triangle (trapèze) vocalique pour représenter les sons étrangers. Il s'agit d'outils traditionnels, qui sont facilement exploitables mais qui peuvent introduire également des erreurs de prononciation car ils créent parfois l'illusion que deux sons, l'un de la LM et l'autre de la LE sont *identiques* puisqu'on utilise le même symbole phonétique pour les retranscrire. Ainsi, pour une meilleure compréhension des caractéristiques acoustiques des sons et du lien articulatoire-acoustique, il est recommandé d'utiliser également des outils comme le visualiseur des spectrogrammes en temps réel ou bien les modèles articulatoires de la production des sons.

4.1.1 Alphabet Phonétique International (API)

La fin du 19^e siècle est marquée par l'apparition de nombreux systèmes de transcription phonétique (Léon, 1993). En France, on utilise l'alphabet des Romanistes ou des Atlas linguistiques et essentiellement l'Alphabet Phonétique International (API) (Léon and Léon, 2007). L'API est proposé en 1886 par des professeurs de langues étrangères, P. Passy, H. Sweet et E. Sievers. Sa création répond aux besoins des enseignants à rendre compte de la prononciation de plusieurs langues étrangères au moyen d'un même alphabet (Léon, 1993). Il s'agit d'un système de notation simple car chaque phonème n'est représenté que par un seul symbole. Grâce à cette invention, la phonétique s'est enfin fait une place dans l'enseignement des langues étrangères (Guimbretière, 1994).

Les principes de la construction de l'API sont décrits dans le livret « The Principles of the International Phonetic Association », publié par l'Association Phonétique Internationale en 1949 (International Phonetic Association, 2010) et ils sont spécifiés pour le français dans l'article de Durand (A paraître). L'alphabet est resté sans grand changement pendant 40 ans (Pullum and Ladusaw, 1996) avant de subir en 1989 une importante révision. D'autres modifications notables sont annoncées en 1993.

Dans sa version d'origine, l'API a été conçu pour représenter les phonèmes de quelques langues, essentiellement le français, l'anglais et l'allemand (Grammont, 1933). Au départ, il ne possédait alors que les symboles de base qui ne permettaient pas de noter des nuances phonétiques. Au fur et à mesure, avec l'apparition de diacritiques, il a été possible de préciser de nombreuses caractéristiques phonétiques, tel que le type de phonation, l'avancement ou rétraction de la langue, l'état du voile du palais ou le rétrécissement secondaire du conduit vocal (Vaissière, 2006). Rochet (1995) souligne l'importance de l'utilisation des diacritiques pour montrer avec précision les différences entre les sons. Ainsi, l'utilisation du symbole [t] sans diacritique ne permet pas de montrer le différent lieu d'articulation de cette consonne en français et en anglais. Alors qu'en français [t̪] est articulé au niveau des dents (c'est une consonne occlusive dentale), en anglais [t̟] s'articule au niveau des alvéoles (c'est donc une consonne occlusive alvéolaire) (Dart, 1998; Delattre, 1953). Mais malgré l'invention de nombreux caractères diacritiques, il est souvent impossible d'exprimer la caractéristique précise du son par un symbole (Vaissière, 2011). C'est en particulier le cas des voyelles.

4.1.2 Triangle vocalique

Un autre outil d'enseignement phonétique proposé au 19^e siècle pour représenter les voyelles d'une langue est un triangle vocalique. Les premiers à établir un triangle vocalique, qui est d'abord de nature articulatoire, sont des professeurs de LE : D. Jones et P. Passy (Delattre, 1968). Il représente les voyelles selon leur aperture, sur l'axe vertical, et leur point d'articulation, sur l'axe horizontal. En 1948, M. Joos convertit le triangle articulatoire en triangle acoustique (Borrell, 1993; Delattre, 1951) avec l'axe vertical représentant le premier formant F1 et l'axe horizontal représentant le deuxième formant F2. La similitude entre les deux triangles montre de façon relativement grossière le lien entre les caractéristiques acoustiques et articulatoires. Dans un premier temps, Joos relie le deuxième formant au déplacement horizontal de la langue et le premier formant au point le plus élevé de la langue. Delattre redéfinit ensuite la relation entre l'aperture et le F1 en précisant que ce dernier est relié à l'ouverture générale du conduit vocal (Delattre, 1951). Les données cinéradiographiques montrent par la suite que la relation entre les mouvements des articulateurs et les effets acoustiques n'est pas aussi simple qu'indiqué au départ dans le triangle vocalique. Ainsi le deuxième formant ne dépend pas uniquement du déplacement horizontal de la langue mais il est dépendant également de l'arrondissement des lèvres (Delattre, 1969). La Figure 28 de Crothers (1978) indique la corrélation entre les articulateurs (constriction du dos de la langue, image 2, position des mâchoires, image 3 et position des lèvres, image 4) et les formants vocaliques F1 et F2 (image 1) correspondants.

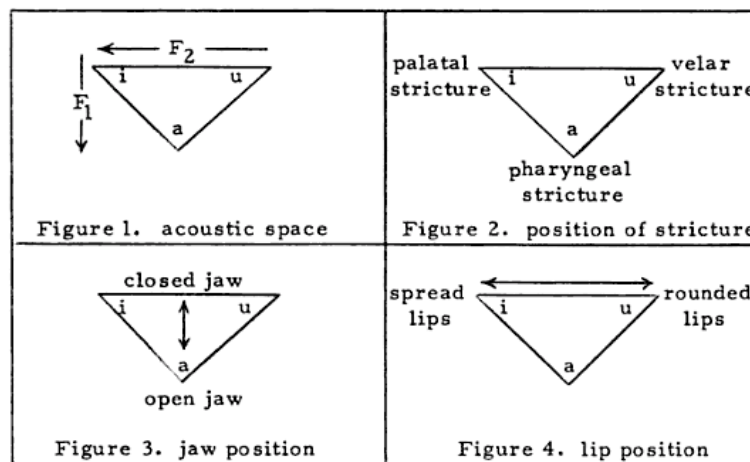


Figure 28 : Corrélation entre triangle acoustique (image 1) et triangles articulatoires (image 2 à 4), selon Crothers (1978)

Dans le cas de voyelles étirées, le deuxième formant dépend du mouvement du dos de la langue et de l'aperture de la mâchoire (Maeda, 1989b). Pour expliquer ces phénomènes compensatoires, d'autres modèles émergent et ils prennent en compte l'action de tous les articulateurs impliqués dans la construction d'une qualité résonancielle, à savoir la langue, les lèvres, la position du larynx et du vélum ainsi que le degré d'ouverture des mandibules (Fant, 1960).

4.1.3 Modèles articulatoires de la production des sons de la parole

Les modèles articulatoires constituent un excellent outil éducatif pour comprendre les lois acoustiques abstraites car les étudiants peuvent facilement comprendre la relation qui réside entre la forme du conduit vocal, difficilement observable autrement, et les effets acoustiques (Arai, 2001).

VTCalcs (Vocal Tract Calculations) est un logiciel permettant de représenter la relation articulatoire-acoustique des sons. Il se base sur le modèle articulatoire de S. Maeda, calculé à partir de

l'analyse de 1000 images du conduit vocal prises au cours de la production de dix phrases françaises par deux locuteurs (Maeda, 1989b). Ce modèle permet de manipuler sept paramètres articulatoires :

- position de la mâchoire : ouverte/ fermée
- position du dos de la langue : avant/ arrière
- forme du dos de la langue : cambrée/ plate
- position de l'apex : relevé/ abaissé
- apertures des lèvres (arrondies/ étirées)
- protrusion des lèvres
- hauteur du larynx

VTDemo - Vocal Tract Acoustics Demonstrator (Huckvale, 2009) est une extension du logiciel VTCalc. Il permet de définir la fréquence fondamentale en choisissant de synthétiser une voix d'homme, de femme ou d'enfant. VTDemo, en plus des paramètres articulatoires de VTCalc, rajoute les paramètres suivants :

- aire de la glotte
- port vélo-pharyngal

L'interface du logiciel est illustrée à la Figure 29 où se trouvent les paramètres articulatoires (à gauche), la forme du conduit vocal (à droite) et les formants des sons résultants (en bas à droite).

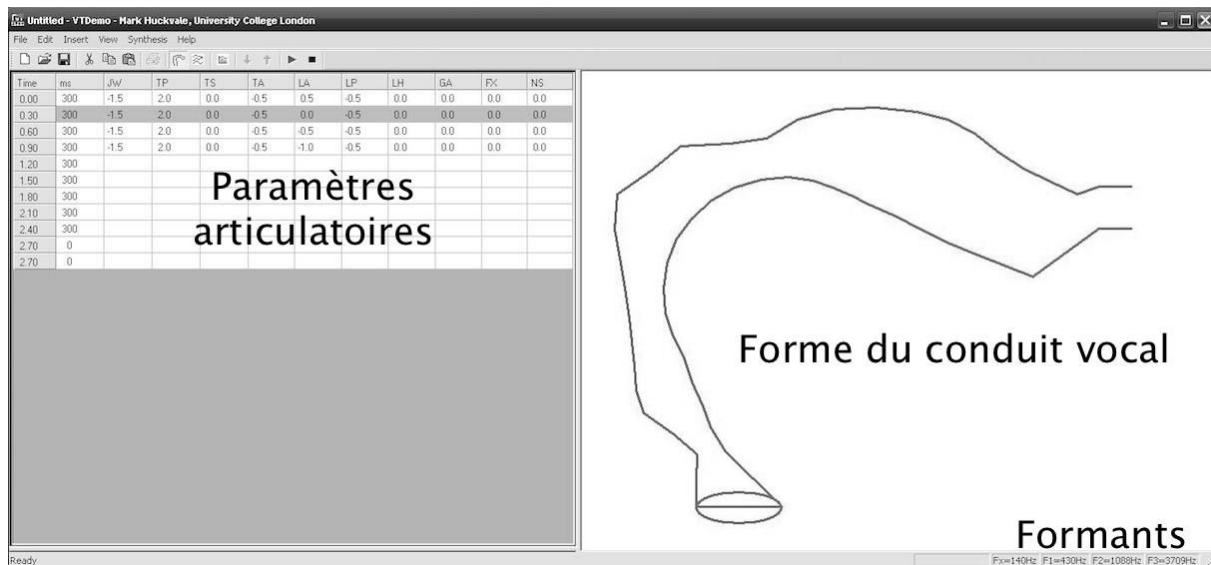


Figure 29 : Interface graphique du logiciel VTDemo (Huckvale, 2009)

Nous avons utilisé ce logiciel pour prédire les effets acoustiques de labialisation, palatalisation et vélarisation phonétique sur les voyelles dans le chapitre 9.

4.1.4 Spectrogramme en temps réel

Le spectrogramme est l'outil indispensable pour l'analyse acoustique de la parole et il représente le temps sur l'axe horizontal, les fréquences sur l'axe vertical et l'intensité des fréquences encodée par le niveau de noir. Le logiciel open source WaveSurfer (www.speech.kth.se/wavesurfer/) a été développé en 2006 par Kåre Sjölander and Jonas Beskow au Centre de technologie de la parole en Suède et il permet de générer des spectrogrammes en temps réel. L'avantage pour les utilisateurs réside alors en visionnage instantané des caractéristiques acoustiques des sons qu'ils sont en train d'émettre. L'enseignant peut alors montrer immédiatement les différentes caractéristiques des sons, ce qui est particulièrement intéressant dans le passage des voyelles françaises [i] à [y] par exemple (voir Figure 30) où l'apprenant peut voir le changement d'affiliation de cavités (dans le cas de [i], la cavité

antérieure détermine essentiellement le troisième formant alors que dans le cas de [y], elle détermine essentiellement le deuxième formant).

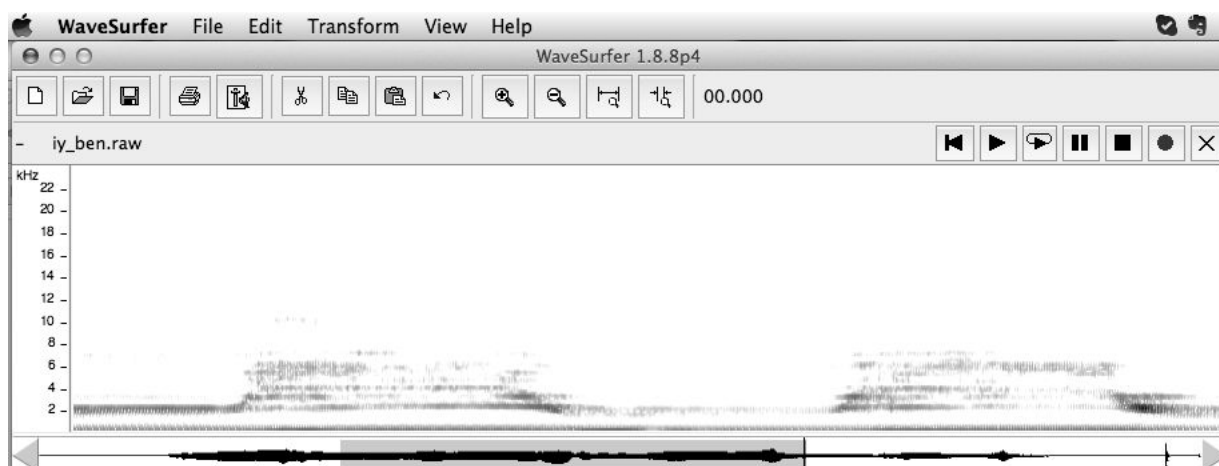


Figure 30 : Interface graphique du logiciel WaveSurfer avec spectrogramme des voyelles [i, y] produites par un Français natif

Les apprenants peuvent également utiliser le logiciel pour vérifier par exemple le caractère focal dans leur production des voyelles françaises [u, o, ɔ, a, y, i], réalisées par les natifs avec un rapprochement et amplification des premier et deuxième formants des voyelles postérieures, des deuxième et troisième formants du [y] et des troisième et quatrième formants du [i].

4.2 Norme phonétique

Les approches communicative et actionnelle, souvent adoptées par les concepteurs des manuels actuels de FLE, soulignent l'importance du travail avec des documents authentiques, qui comportent les variations de la langue française (Puren, 2006). Mais alors face à ces variétés, laquelle doit être choisie comme étant la norme ?

Les apprenants souhaitent en général adopter une variante neutre, non marquée (Detey *et al.*, 2010; Wachs, 2011). Ils veulent pour la plupart acquérir un accent qui passe partout dans l'objectif de parler sans se faire remarquer (Billières and Borrell, 1990).

Or la norme en phonétique est difficile à définir (Borrell and Billières, 1989) et les auteurs des manuels du FLE ne définissent en effet jamais le modèle qui est sous-jacent dans les enregistrements accompagnant ces manuels (Billières and Borrell, 1990). Elle est pourtant nécessaire, comme le précise Flege (1987b, p. 288) : « *Instrumental studies are based ... on the assumption that a phonetic norm exists for the L2* ». Mais faut-il la dégager à partir de la parole de laboratoire, hautement contrôlée, ou bien à partir de la parole spontanée ? Valdman (2000) stipule que la norme doit refléter l'usage réel de la langue. Les variantes retenues pour définir la norme doivent être les plus fréquentes et les plus valorisées par les natifs. Detey and Racine (2012), eux, font la différence entre une norme d'usage qui s'applique dans la vie quotidienne et une norme de prestige. Detey *et al.* (2010) définissent le « français de référence » perceptivement. Ils font entendre à 58 auditeurs français natifs, venant de la ville de Paris, Marseille et Toulouse, des extraits de parole de cinq locuteurs différents dont deux enseignants de FLE, l'un de la région parisienne, et l'autre originaire du sud de la France, puis un troisième locuteur parisien, un quatrième venant de Biarritz et enfin un cinquième locuteur venant de l'Ile de la Réunion. Les auditeurs doivent indiquer sur une échelle de 1 à 6 si la prononciation des séquences entendues est proche du français standard (6) ou pas du tout (1). Le résultat montre que les enseignants de FLE (l'un originaire de l'Ile-de-France, l'autre du sud de la France) et le locuteur parisien sont considérés comme étant plus proches du « français de référence » que les personnes qui viennent de Biarritz et de la Réunion et qui ne sont pas enseignants. Ceci prouve que la norme est bien

présente dans l'esprit des auditeurs (même si sa définition pose de nombreux problèmes) et que par conséquent il convient de la conserver dans l'enseignement des langues.

Au cours du 20^e siècle, la norme du français a évolué (Billières and Borrell, 1990). Avant 1960, le bon français est celui de la haute bourgeoisie de Paris. On le dénomme alors le « français parisien cultivé » et il est décrit surtout dans les ouvrages de M. Grammont et P. Fouché. A partir de 1960, un nouveau modèle émerge, il s'agit du « français standard ». Le « bon français » n'est plus uniquement celui de Paris mais également celui des autres grandes villes de France. Il s'agit du français des médias qui rayonne sur la France entière.

Après 1980, le modèle de référence est le « français standardisé » qui caractérise le parler de la classe moyenne favorisée, sans marque régionale particulière. C'est le français de tout l'Hexagone. Désormais, le français normé concerne de plus en plus de monde. Cette ouverture est le résultat de nombreuses enquêtes phonologiques et phonétiques, initiées en 1945 par Martinet (1945) qui montre que le français parisien n'est qu'une des variantes du français parlé. Les enquêtes se poursuivent essentiellement dans les années 1960-70. En 1973, Martinet and Walter (1973) publient le « Dictionnaire de la prononciation française dans son usage réel » et P. Léon élabore par la suite un manuel de la prononciation du français standard en prenant en compte les résultats des enquêtes menées (Léon, 1976).

Dans les sciences phonétiques, le français standard est décrit, selon les auteurs, comme le français du Nord (De Mareüil *et al.*, 2010), le français non-méridional (Pustka, 2011) ou encore le français parisien (Miller *et al.*, 2011). Selon ces derniers, « *Parisian French (standard French) is spoken in and around Paris as well as in many other parts of France (excluding the south)* ».

Parmi les termes utilisés en phonétique pour décrire le français standard, nous avons choisi celui du « français non-méridional ». Ainsi, les locuteurs qui constituent la « référence » dans cette thèse sont des francophones natifs non-méridionaux.

4.3 Méthodes d'enseignement des langues étrangères

Une méthodologie renvoie à un ensemble de procédés et manières de faire cohérents utilisés dans un programme d'enseignement (Billières and Spanghero-Gaillard, 2005). La cohérence de chaque méthodologie repose sur un noyau dur qui est constitué d'un ensemble de méthodes privilégiées, telles que la méthode active (participation active de l'apprenant) ou transmissive (l'enseignant transmet les savoirs, l'apprenant y est attentif), orale (l'apprenant parle) ou écrite, directe (utilisation de la L2) ou indirecte, inductive (établissement de règles à partir de la production) ou déductive (exposition des règles d'abord), conceptualisatrice (l'apprenant appréhende les règles) ou imitative (l'apprenant répète de manière intense pour acquérir des automatismes) (Puren, 1988; Puren, 2001).

Le terme « méthode » peut-être employé soit dans le sens de « méthodologie » (pour parler de la méthode directe, audio-orale, etc.), soit dans les sens de matériel didactique, comme « la méthode *Panormama* » ou « la méthode *Assimil* » qui est constitué de manuels, (avec souvent un livre du maître, guide pédagogique et cahier d'exercice) et des CD ou DVD, construits selon une planification didactique (la présentation du contenu) et pédagogique (les objectifs de l'apprentissage).

En méthodologie, on distingue les méthodes conventionnelles et non-conventionnelles (Billières and Spanghero-Gaillard, 2005). Parmi les méthodes conventionnelles peuvent être classées la méthode traditionnelle, directe, audio-orale ou phonétique intégrale, SGAV, l'approche communicative et actionnelle. Notons que le terme « approche » désigne ici une méthodologie qui se veut souple, ouverte et non-dogmatique. Les méthodes non-conventionnelles sont représentées par les

méthodes silencieuses où la responsabilité revient à l'apprenant (*Silent Way*), la méthode *OK-feelings* où l'on vise le bien-être, la suggestopédie qui suscite également la relaxation, la méthode communautaire visant la dynamique du groupe des apprenants, etc. (Puren, 1990). Elles ont été élaborées aux Etats-Unis essentiellement par des psychologues et elles sollicitent la créativité et l'imagination de l'apprenant pour garantir la réussite (Constantinescu-Stefanel, 2006). Contrairement aux méthodes conventionnelles, elles ne sont fondées sur aucun noyau dur et théorie de référence sous-jacente (description linguistique ou la psychologie de l'apprentissage) et donc ne peuvent pas vraiment détenir le statut de « méthodologie ».

Dans ce qui suit, seules les méthodes conventionnelles sont détaillées.

4.3.1 Méthodes conventionnelles

Méthode traditionnelle : La première méthode utilisée dans l'enseignement des langues est la méthode traditionnelle, appelée également « méthode grammaire-traduction ». Elle est utilisée à partir du 16^e siècle, d'abord pour enseigner le latin et le grec, puis pour enseigner les langues vivantes (Besse, 1985; Stoean, 2006). Son objectif premier est de donner l'accès à des textes littéraires et la compétence écrite est donc au centre de cette méthodologie (CREDIF, 1972). La prononciation est abordée à travers des dictées et par apprentissage par cœur des morceaux cités. Il s'agit d'une méthode transmissive (et donc passive), indirecte, déductive et répétitive.

Méthode directe : La méthode directe a émergé en Europe au début du 20^e siècle grâce à, entre autre, P. Passy, H. Sweet ou H. E. Palmer (Besse, 1985; Dobre, 2006b). Il s'agit d'une méthode active, inductive et directe où seule l'utilisation de la L2 est permise dès la première leçon. L'enseignant a recours à des gestes, des mimiques et des dessins. On sollicite l'écoute, la vue et les mouvements du corps entier. L'oral y occupe une place importante et les transcriptions phonétiques sont systématiques dans presque tous les manuels.

Méthode audio-orale : La méthode audio-orale émerge en Amérique du Nord en 1950 (Besse, 1985; Dobre, 2006a). Elle jouit alors d'un grand succès et se propage en France à partir de 1965. A la base, elle s'appuie essentiellement sur les exercices structuraux des linguistes L. Bloomfield, E. Sapir et F. Boas ainsi que sur une expérience didactique conduite par l'Armée américaine au cours de la deuxième guerre mondiale. Les enseignants-linguistes sont des locuteurs natifs de la L2 et les apprenants reproduisent de manière intensive les modèles de la langue orale. Cette méthodologie repose ainsi sur la méthode imitative et orale (quoique non intégrale car l'apprenant est amené à faire des transcriptions écrites des dialogues). La méthode de l'Armée américaine avec un travail intensif en laboratoires de langues a montré qu'il était possible d'apprendre à parler une L2 très rapidement.

Les linguistes américains C.C. Fries et R. Lado, adoptent cette méthode, très intensive, d'enseignement des LE à un public plus large. Ils travaillent désormais avec des enregistrements faits par des locuteurs natifs et la progression se fait à l'aide d'exercices structuraux qui permettent de créer des automatismes comportementaux. Ainsi les théories de référence sont le structuralisme et le behaviorisme. Les élèves commencent par des structures faciles, qui sont supposées entraîner le moins d'interférences possibles. La progression est ainsi programmée selon la difficulté prédite par l'analyse contrastive, introduite par Lado (1957).

Méthode phonétique intégrale : En même temps, Delattre (1944) élabore la méthode phonétique intégrale destinée aux écoles de langues et aux universités. Elle ne vise plus la quantité et la rapidité, comme la méthode de l'Armée mais la qualité. Selon cette méthode, un cours idéal de LE, au départ exclusivement oral, devrait s'appuyer sur des matériaux méticuleusement construits et présenter les difficultés phonétiques graduellement, c'est-à-dire aborder d'abord la prosodie et les principes généraux d'articulation (antériorité, position de la langue, des lèvres, la loi de position), puis les sons individuels, mais toujours intégrés dans des unités de sens. Un cours de français pour débutants devrait ainsi éviter utilisation de tout support écrit (d'où la dénomination « méthode orale

‘intégrale’ »), afin d’empêcher l’association d’un son à une image visuelle (graphème, symbole phonétique) mais favoriser au contraire l’association directe du son au sens. Seulement quand l’apprenant maîtrise authentiquement un mot ou une phrase, il peut en apprendre son orthographe. Delattre (1947) élabore un manuel d’introduction des habitudes du français oral basé sur la méthode intégrale.

Approche communicative : L’enthousiasme induit par le travail en laboratoires de langues diminue et les années 1970 marquent l’apparition de l’approche communicative qui introduit le concept de « compétence » (Billières and Borrell, 1990). L’apprenant se trouve désormais au centre de l’apprentissage et l’objectif principal n’est plus une prononciation native mais plutôt la capacité à communiquer et à se faire comprendre dans les situations de la vie quotidienne (Puren, 2006). Ainsi, les exercices de prononciation sont moins nombreux et l’enseignement de la phonétique est peu à peu délaissé (Billières and Borrell, 1990; Derwing and Munro, 2009; Leather, 1983; Wachs, 2011). Il s’agit d’une méthode active et conceptualisatrice. Elle est en même temps inductive (l’apprenant observe les régularités de la L2 à partir des productions) et déductive (il teste par la suite les règles observées).

Le Tableau 22 récapitule les méthodologies conventionnelles en indiquant la période de leur apparition, l’objectif général, le noyau dur les constituant, les théories sous-jacentes et le support de l’enseignant.

Méthodologie	Période	Objectif général	Noyau dur	Théories sous-jacentes	Support de l’enseignant
Méthode traditionnelle	A partir du 16 ^e siècle	Faciliter l’accès aux textes classiques	Méthode transmissive, indirecte, déductive, répétitive	Pas de théorie précise	Textes littéraires
Méthode directe	Début 20 ^e	Apprendre activement une L2, parler	Méthode directe, inductive, active	Empirisme (Passy)	Gestes, mimiques, dessins, API
Méthode audio-orale	1940 1965 en France	Comprendre et parler, puis lire et écrire	Méthode orale, imitative, répétitive	Linguistique structurale et behaviorisme (Bloomfield, Harris, Lado, Skinner)	Magnétoscope à bande, analyse contrastive
Méthode phonétique intégrale	1940	Comprendre et parler comme un natif	Méthode orale intégrale, imitative, répétitive	Linguistique structurale et behaviorisme	Magnétoscope à bande Banni : écrit (chez les débutants)
Approche communicative	1980	Parler et communiquer dans les situations de la vie courante sans but d’avoir une performance native	Méthode conceptualisatrice, active, inductive et déductive	Linguistique pragmatique (Hymes, Austin, Searle), psychologie cognitive	Documents authentiques avec nombreuses variétés de la L2

Tableau 22 : Méthodologies conventionnelles d’enseignement des langues étrangères

4.3.2 Principes de correction phonétique

Les exercices en phonétique corrective sont construits essentiellement à partir de deux méthodes : la méthode verbo-tonale d’intégration phonétique (MVT) et la méthode articulatoire (MA). La MA a apparu bien avant la MVT et elle est largement utilisée. En quoi les deux méthodes diffèrent ?

Méthode verbo-tonale d’intégration phonétique : La MVT fait partie de la méthodologie SGAV (structuro-globale audio-visuelle, émergeant dans les années cinquante) et elle est le fruit d’une

collaboration des chercheurs de Saint-Cloud, sous la direction de G. Gougenheim et P. Rivenc (travaillant avec des malentendants), et de l'Université de Zagreb, représentés par P. Guberina (travaillant avec des apprenants croates). Elle se base sur le concept du « crible phonologique » (Intravaia, 1993; Magnen *et al.*, 2005; Murillo, 1982), développé d'abord par Polivanov (1931) et repris par (Troubetzkoy, 1957): les difficultés de prononciation sont dues à des problèmes de perception qu'il faut rééduquer en premier lieu (Guberina *et al.*, 1965; Landercy, 1976; Magnen *et al.*, 2005; Urbain, 1975). Dans la rééducation, on fait appel aux mouvements du corps entier, de la mélodie et du rythme (Guberina, 1970). On introduit l'usage de l'appareil *Suvag* qui permet d'éliminer de nombreuses fréquences non perçues par l'apprenant (Guberina, 1973) et de remodeler ainsi l'oreille (Cureau, 1973). On considère qu'il est inutile de travailler la prononciation tant que la perception de l'apprenant continue à appliquer le filtre de la langue maternelle (Lebel, 1969). Dans les exercices de prononciation, on se sert des entourages facilitants pour favoriser la bonne prononciation des sons (le timbre vocalique peut être assombri ou éclairci au contact d'une consonne). L'accent est porté sur la communication dans des situations de la vie courante (linguistique de la parole). En revanche la méthode verbo-tonale ne recourt ni à l'explication articulatoire, ni à l'analyse contrastive, ni à l'API (Intravaia, 1976).

Méthode articulatoire : La MA se repose sur la description articulatoire d'un son que l'apprenant d'une L2 doit connaître. L'articulation est au centre de l'apprentissage et les facteurs relevant de l'audition sont négligés. Le raisonnement est alors inverse de celui de la MVT. Les enseignants ont recours à des schémas articulatoires des cibles vocaliques et consonantiques (coupes sagittales, lèvres vues de face, etc.), à des descriptions de la position exacte des organes de la parole dans la production des sons et à l'analyse contrastive entre la L1 et la L2 en termes de base articulatoire. L'apprenant est invité à reproduire des sons en étant adossé au dossier de la chaise et avec le buste relevé. Le mouvement est donc absent, contrairement à la MVT. Les exercices commencent souvent par la prononciation de sons isolés, puis des logatomes, ensuite des mots isolés et enfin des phrases isolées. L'influence de la qualité d'un son sur le son adjacent n'est pas prise en compte et les exercices avec des contextes facilitants ne font donc pas partie du programme.

Méthode des oppositions phonologiques : Enfin, une troisième méthode est largement répandue dans les manuels de LE, à savoir la méthode des oppositions phonologiques. En s'appuyant sur la linguistique structurale, l'acquisition des phonèmes se fait à partir des unités de sens. L'avantage est que l'apprenant comprend tout de suite l'utilité de réaliser correctement la qualité d'un son (Abry and Veldeman-Abry, 2007).

Ces méthodes, avec parfois la correction fondée sur la transcription phonétique, se trouvent à la base d'exercices et d'activités en matière de correction phonétique que l'on trouve dans de nombreux manuels et sur des sites web dédiés à l'apprentissage des langues étrangères.

4.4 *Evaluation du niveau phonétique*

Le niveau phonétique exact de l'apprenant est difficile à évaluer. Le Cadre européen commun de référence pour les langues CECRI (Conseil de l'Europe, 2001) permet de définir les compétences et les niveaux de langues européennes. Le Cadre est le résultat d'une dizaine d'années de travail, initié par le Conseil de l'Europe et réalisé par de nombreux spécialistes de langues européennes. Il s'inscrit dans l'approche actionnelle, qui place l'apprenant au centre du processus d'apprentissage. L'apprenant est alors considéré comme un acteur social qui doit accomplir un certain nombre de tâches. Le Cadre définit trois niveaux d'utilisateurs : élémentaire (A), indépendant (B) et expérimenté (C). Ces niveaux sont de plus subdivisés pour proposer une définition plus fine : niveau introductif (A1), de survie (A2), seuil (B1), avancé (B2), autonome (C1), de maîtrise (C2). Les différentes compétences évaluées sont la compréhension et l'expression, écrite et orale. Les examinateurs, afin de déterminer le niveau de langue, doivent remplir des grilles d'évaluation.

PREMIERE PARTIE : INTRODUCTION

Le Cadre définit la compétence phonologique et orthoépique. La compétence orthoépique exige la connaissance des conventions orthographiques et des signes de ponctuation ce qui se traduira par l'utilisation de l'intonation et du rythme approprié.

Concernant la maîtrise du système phonologique, les descripteurs formulent les niveaux comme suit (p.92) :

A1 : La prononciation d'un répertoire très limité d'expressions et de mots mémorisés est compréhensible avec quelque effort pour un locuteur natif habitué aux locuteurs du groupe linguistique de l'apprenant/utilisateur.

A2 : La prononciation est en général suffisamment claire pour être comprise malgré un net accent étranger mais l'interlocuteur devra parfois faire répéter.

B1 : La prononciation est clairement intelligible même si un accent étranger est quelquefois perceptible et même si des erreurs de prononciation proviennent occasionnellement.

B2 : A acquis une prononciation et une intonation claires et naturelles.

C1 : Peut varier l'intonation et placer l'accent phrastique correctement afin d'exprimer de fines nuances de sens.

C2 : Comme C1.

Il est évident que les critères pour définir les différents niveaux en phonétique au niveau du CECRI sont flous (Harmegnies *et al.*, 2005). Notons que les niveaux C1 et C2 ne sont pas distingués.

Ainsi, il semble difficile d'évaluer le niveau phonétique des apprenants. Cunningham-Andersson and Engstrand (1989) proposent de corrélérer l'accent étranger perçu au nombre d'erreurs de prononciation. Les apprenants qui font plus d'erreurs phonétiques seraient ainsi jugés comme ayant un accent étranger plus fort. Selon les auteurs, il serait donc possible d'évaluer l'accent étranger à partir d'un petit extrait lu, et par conséquent de regrouper les apprenants selon leur niveau phonétique. Toutefois le type de faute commise est également important car certains écarts font penser plutôt à des variantes régionales.

5. Accent étranger

L'accent du pays où l'on est né demeure dans l'esprit et dans le cœur, comme dans le langage.

François de la Rochefoucauld

Résumé : Le **cinquième chapitre** traite de « l'accent étranger » qui peut être défini comme une déviation par rapport à la norme (Clarke and Garrett, 2004) et qui concerne aussi bien la production que la perception (McAllister, 1998). Les facteurs responsables de l'accent étranger sont nombreux (pour une revue de littérature, voir par exemple Flege (1988b) : âge au début de l'apprentissage, qualité et quantité d'input, graphie, utilisation de la L2 par rapport à la L1, nombre d'années de résidence dans le pays où est parlée la L2, genre, type d'apprentissage (formelle avec supports écrits ou naturelle en immersion), personnalité et motivation de l'apprenant, origine linguistique et culturelle de l'auditeur natif (les monolingues seraient moins tolérants que les personnes parlant plusieurs LM/LE) et propriétés phonétiques des sons. En effet, selon Flege and Hillenbrand (1984), les sons de la LE *similaires* aux sons maternels sont produits avec moins de précision que les sons *nouveaux* et *identiques*. Le dernier facteur nécessite la délimitation des frontières entre les sons *similaires*, *nouveaux* et *identiques* qui devrait être systématiquement basée, selon Rochet (1995), sur la comparaison des symboles phonétiques (1), des caractéristiques acoustiques (2) et des caractéristiques perceptives (3). Enfin, les conséquences de l'accent étranger peuvent être négatives (baisse de l'intelligibilité, discrimination des locuteurs non-natifs, irritation des interlocuteurs) ou positives (accent chic, noble ou sexy).

L'accent étranger peut être défini comme une déviation par rapport aux prototypes des locuteurs natifs (Clarke and Garrett, 2004). Il est ainsi souvent associé à la notion de différence (Derwing and Munro, 2009). Pour Scovel (1978), il s'agit d'un signe phonologique qui montre que le sujet parlant n'est pas un locuteur natif. Flege and Hillenbrand (1984) expliquent l'accent étranger, au moins en partie, non seulement par les différences phonologiques, mais aussi phonétiques qui existent entre la L1 et L2. Selon Harmegnies *et al.* (2005), l'accent étranger caractérise le locuteur et il ne peut être modifié que par apprentissage. Grosjean (2013) avertit que l'accent étranger est un phénomène courant même en situation de bilinguisme et qu'il est rare de s'en défaire.

La plupart des auteurs utilisent la notion de l'accent étranger pour faire apparaître une différence en production. Ainsi, Rochet (1995) stipule que l'accent étranger est le résultat d'une reproduction erronée des sons de la LE. Or, c'est un phénomène qui touche également la perception de l'apprenant même s'il n'est pas possible de l'observer directement (McAllister, 1998; McAllister, 2000).

L'accent étranger serait plus rapidement détecté au niveau des phrases plutôt que des mots isolés. Major (1987) distingue ainsi entre l'accent global, perçu au niveau de la phrase, et local, perçu au niveau du mot. Il a effectué une étude avec 53 étudiants brésiliens et sept Américains natifs qui ont produit en anglais une liste de mots, une liste de phrases et un texte court. Dix Américains natifs ont ensuite évalué la force de l'accent étranger au niveau global (sur une échelle de 1 à 256), puis ont effectué un test d'identification des suites [ɛt] et [æt] en indiquant sur une échelle de 1 à 5 la force de l'accent étranger. Le résultat montre que selon les anglophones natifs, les 53 Brésiliens avaient produit les mots isolés plus authentiquement que les phrases entières alors que le résultat était inverse pour les productions des locuteurs américains natifs.

A l'inverse, van Wijngaarden *et al.* (2002) observent que les sons isolés produits par les non-natifs peuvent être moins bien discriminés que lorsqu'ils apparaissent en contexte. Dans la parole, leur intelligibilité augmente car le contexte joue un rôle puissant.

Southwood and Flege (1999) constatent que l'accent étranger peut être détecté rapidement. Flege (1984) montre que l'accent étranger peut être perçu au niveau de la syllabe, d'un phone ou même dans une portion de 30 ms du burst d'une occlusive.

De nombreux chercheurs s'accordent à dire que l'apprentissage de la L2 diffère de celui de la LM (Flege, 1991). Les apprenants tardifs d'une LE parlent souvent avec « l'accent étranger » (Derwing and Munro, 2009). Mais quels sont les facteurs qui agissent sur le degré de l'accent étranger ? Parmi les variables les plus influentes sont cités l'âge au début de l'apprentissage, la qualité et la quantité d'input et l'utilisation relative de la L2 par rapport à la L1 (Piske *et al.*, 2001). Les autres variables à prendre en compte sont la durée du séjour dans le pays où est parlée la L2, les facteurs affectifs, le sexe du locuteur et les facteurs linguistiques. Le degré de l'accent étranger dépend également de la perception de l'auditeur natif.

5.1 Age au début de l'apprentissage

L'âge au début de l'apprentissage d'une LE joue un rôle considérable dans la production et la perception authentique de l'apprenant. De nombreux chercheurs mettent en évidence que plus tôt on commence à apprendre une LE dans la vie, plus il y a de chance à acquérir une prononciation native (Flege, 1988a; Grosjean, 2013; Piske *et al.*, 2001).

Dans certaines études antérieures (Krashen, 1975; Lenneberg, 1967; Scovel, 1978), l'âge était considéré comme un obstacle biologique qui rendait l'apprentissage authentique impossible chez les adultes. Les études plus récentes (Flege, 1991; Oyama, 1976; Piske *et al.*, 2001) montrent que la capacité d'apprentissage d'une langue reste intacte durant toute la vie. L'apprentissage de la LE est étroitement lié à l'acquisition de la LM, plutôt qu'à l'âge comme tel (Iverson *et al.*, 2001). Si on apprend une LE alors que les catégories phonémiques de la LM ne sont pas encore complètement figées, il sera plus facile de modifier le système phonético-phonologique de la LM pour inclure les catégories de la LE (Flege, 1988a; Flege, 1995). Oyama (1976) remplace le terme « période critique », précédemment utilisé pour décrire un phénomène biologique, par « période sensible ». Elle a conduit une étude avec 60 immigrants italiens entre 6 et 20 ans, qui ont commencé à apprendre l'anglais à leur arrivée aux Etats-Unis. L'âge des immigrants s'est révélé comme étant un facteur agissant sur la perception de l'accent étranger par les anglophones natifs. Selon les Américains natifs, les immigrants qui sont arrivés aux Etats-Unis plus tôt, et donc ont appris l'anglais plus jeunes, parlaient avec moins d'accent que ceux qui sont arrivés à un âge plus avancé.

Tahta *et al.* (1981) ont réalisé une étude en Grande Bretagne avec 109 immigrants, venant de pays différents qui ont commencé à apprendre l'anglais à leur arrivée, entre 6 et 77 ans. Dans une tâche de production, ils ont lu un paragraphe d'un texte en anglais et la force de l'accent global a été jugé par la suite par trois auditeurs anglais natifs sur une échelle de 0 (pas d'accent étranger) à 2 (accent étranger très prononcé). Là encore, l'âge s'est avéré comme étant un facteur important sur la perception du degré de l'accent étranger par des natifs de l'anglais car les immigrants qui avaient commencé l'apprentissage à l'âge de 6 ans ne présentaient pas, selon les informateurs natifs, d'accent étranger. Ainsi les auteurs situent la période sensible pour l'apprentissage authentique d'une LE aux alentours de 6 ans. Si l'apprentissage débute avant cet âge, il n'y a pas de transfert négatif de la LM sur la LE. Pour Flege (1991), l'âge du début de l'apprentissage de la LE influence principalement l'acquisition des sons *similaires* qui peuvent être maîtrisés à condition que l'apprentissage commence dans la petite enfance, avant 5 à 7 ans. Au-delà de cette période sensible, la production de ces sons ne pourra plus être authentique.

A l'inverse, Bongaerts *et al.* (1997) ne soutiennent pas l'idée de la période sensible. Dans leur étude avec 10 apprenants de l'anglais de LM néerlandaise, ils démontrent la possibilité d'acquérir une prononciation authentique même si l'apprentissage a commencé après l'âge de 12 ans, conclusion à laquelle parviennent également Palmen *et al.* (1997) dans une étude similaire.

De nombreuses études montrent que l'effet de l'âge ne peut être considéré isolément car les facteurs agissant sur l'accent étranger sont nombreux (Leather and James, 1991).

5.1.1 Période critique

La théorie développée par Lenneberg (1967) postule l'existence d'un âge critique - aux alentours de 12 ans - pour l'apprentissage d'une langue, qu'elle soit maternelle ou étrangère qui s'expliquerait au niveau neuronal. Ainsi à l'adolescence, le système nerveux est réorganisé et le cerveau subit une latéralisation et par conséquent perd en plasticité. Lenneberg (1967) se base sur les données cliniques, telles que l'endommagement du cerveau, suite à un accident pour expliquer l'effet de l'âge sur l'apprentissage d'une langue et ses découvertes concernent essentiellement le domaine de la syntaxe ou de la sémantique. Scovel (1978) s'intéresse, lui, directement à l'acquisition de la phonologie de la LE.

Selon l'hypothèse de la période critique, il est impossible d'acquérir une LE à l'âge adulte sans accent étranger. Il serait donc inutile d'essayer de changer la production et perception d'un apprenant «tardif» par un apprentissage quelconque, car il serait biologiquement impossible de les modifier.

Ce concept trouve son parallèle dans les études sur la période de cristallisation des oiseaux. Les oiseaux chanteurs apprennent à chanter, tout comme l'homme apprend à parler (Brainard and Doupe, 2002). Pour que l'apprentissage ait lieu, les oiseaux ont besoin de contact avec leur tuteur. Les chants doivent être acquis pendant la période de cristallisation qui est composée de la phase sensorielle, suivie de la phase sensori-motrice. Durant la phase sensorielle, les oiseaux écoutent les tuteurs chanter pour se faire des représentations mentales de ces chants. Pendant la phase sensori-motrice, ils s'exercent eux-mêmes. Cette phase est très importante pour l'apprentissage. Si les oiseaux deviennent sourds alors qu'ils n'ont pas pu s'entraîner et entendre leur propre chant, ils produiront des chants anormaux (Leather, 1983). La période de cristallisation est donc parfois comparée à la période critique pour l'apprentissage d'une langue par un homme.

Flege (1987a) pour sa part présente des arguments contre la période critique pour un apprentissage linguistique. Selon lui, la période critique s'applique globalement à toute une espèce, plutôt qu'à des individus et elle concerne uniquement les comportements qui ne peuvent être ni oubliés ni modifiés. Or l'apprentissage phonétique est possible même à l'âge adulte ; de nombreuses recherches sur le bienfait des entraînements phonétiques en témoignent (Catford and Pisoni, 1970; Leather, 1983; Rochet, 1995).

5.1.2 Crible et surdité phonologique

Une autre hypothèse reliée à l'âge du début de l'apprentissage est celle du crible et de la surdité phonologique du Cercle Linguistique de Prague. Polivanov (1931) utilise la notion du crible pour décrire la perception des sons étrangers. Selon l'auteur, tous les sons passent par le filtre phonologique de la langue maternelle, établi dans la petite enfance, et sont interprétés en fonction de cette dernière. Trubetzkoy (1939) reprend et développe ce concept en introduisant la notion de surdité phonologique qui stipule que les locuteurs non-natifs sont véritablement sourds aux sons qui n'ont pas un rôle fonctionnel dans leur LM. La surdité phonologique est également le concept-clé de la méthode SGAV qui propose une correction de l'audition par des fréquences optimales (Landeracy and Renard, 1976).

5.2 Qualité et quantité d'input

A propos de l'influence de la qualité et quantité d'input linguistique sur l'apprentissage, Grosjean (2013) a écrit : « *You are what you eat... phonetically.* » Cette citation illustre l'idée que l'effet de la qualité et quantité d'input phonétique est important (Bohn and Flege, 1997).

Escudero (2001) a effectué une étude avec deux groupes de 30 apprenants espagnols de l'anglais. Un groupe apprenait l'anglais de l'Ecosse, qui s'appuie essentiellement sur des caractéristiques spectrales pour distinguer entre les voyelles i/ɪ, alors que l'autre groupe apprenait l'anglais de l'Angleterre du sud, qui utilise plutôt l'indice de durée pour différencier les deux voyelles i:/i. Ainsi, les deux groupes d'apprenants espagnols, exposés à des variétés différentes de l'anglais, ont assimilé les voyelles différemment. Le premier groupe les a assimilées à deux catégories vocaliques espagnoles [i], [e] alors que le deuxième les a assimilées à une seule catégorie [i]. Cette expérience démontre l'importance de l'input dans l'apprentissage d'une langue.

5.3 Utilisation relative de la L2 par rapport à la L1

La variable d'utilisation relative de la L2 par rapport à la L1 n'est prise en considération que depuis peu de temps (Piske *et al.*, 2001). Son effet est observé par exemple par Tahta *et al.* (1981) chez 109 immigrants de LM différentes ayant appris l'anglais et par Piske *et al.* (2001) chez des locuteurs italiens parlant l'anglais. Ces derniers ont examiné l'accent étranger chez 43 hommes et 47 femmes bilingues anglo-italiens, qui ont commencé l'apprentissage de l'anglais durant leur enfance ou à l'âge adulte. Leurs productions (lecture de phrases et de mots, description d'une image, narration spontanée) ont été évaluées par neuf anglophones natifs sur une échelle de un (accent étranger très prononcé) à neuf (pas d'accent étranger). Le résultat montre que les bilingues qui utilisent leur L1 plus fréquemment gardent un accent étranger plus prononcé dans la L2. Ce résultat concerne les deux groupes de bilingues.

5.4 Nombre d'années de résidence dans le pays où est parlée la L2

Tahta *et al.* (1981) n'observent pas d'effet de la variable du nombre d'années de résidence dans le pays où est parlée la L2 sur le degré de l'accent étranger. Mais comme leur expérience ne prend en compte que les sujets ayant séjourné dans le pays depuis minimum 2 ans, il est possible que l'effet de cette variable soit observable dans les premiers mois suivant l'arrivée. Flege (1988a) le confirme en constatant un progrès initial rapide, immédiatement après l'arrivée dans le nouveau pays, suivi d'une stagnation. Toutefois les études ayant reporté l'effet de cette variable sur l'accent étranger ont montré qu'il était néanmoins moins important que celui de l'âge au début de l'apprentissage (Piske *et al.*, 2001).

5.5 Genre

L'effet du genre sur la force de l'accent étranger perçu est discutable : alors que certaines études ont montré que les femmes seraient plus performantes que les hommes dans l'acquisition authentique de la LE, d'autres ne supportent pas cette idée (pour revue de littérature voir l'article de Piske *et al.* (2001). Par exemple Asher and Garcia (1969) ont effectué une étude avec 71 immigrants cubains qui sont arrivés aux Etats-Unis entre 7 et 19 ans. Ils ont trouvé que l'acquisition quasi native de l'anglais est apparue chez plus de filles que de garçons.

Tahta *et al.* (1981) ont également montré que les filles et adolescentes sont plus performantes que les jeunes garçons.

5.6 Types d'apprentissage

De nombreuses études ont montré un effet positif de l'entraînement phonétique sur l'intelligibilité et la compréhensibilité de la LE (Flege, 2003). Certaines méthodes se focalisent sur les sons, d'autres méthodes privilégient le travail sur la prosodie. Cependant, Derwing *et al.* (1998) stipulent que seule la méthode globale, qui travaille les sons et la prosodie, est efficace.

Rochet (1995) observe le bienfait d'un entraînement perceptif chez 12 Chinois mandarin, habitant au Canada depuis 2 à 4 ans, à percevoir le contraste entre les consonnes occlusives sourdes et sonores (labiales, dentales et vélaires) du français. Il observe que même si l'on se contente d'entraîner la perception des consonnes labiales et vélaires, l'amélioration est notable en ce qui concerne les consonnes coronales. En revanche, si l'on entraîne la perception des consonnes dans une seule position, par exemple initiale, il est impossible de généraliser le progrès dans les autres positions, par exemple médiane ou finale. Ceci s'explique par le fait que l'actualisation phonétique du voisement de la consonne dépend de sa position.

L'apprentissage d'une LE peut se faire d'une manière naturelle, comme c'est le cas des enfants qui apprennent une nouvelle langue en immersion, ou d'une manière formelle (Asher and Garcia, 1969). Dans l'apprentissage formel, on utilise d'habitude des documents écrits et la présentation de l'écrit peut avoir des effets négatifs sur un apprentissage phonétique (Flege, 1987c; Flege, 1988b).

5.6.1 Ecrit

L'écrit, selon les langues, peut établir une relation entre la graphie et la sonie plus au moins régulière (Harmegnies *et al.*, 2005). Le tchèque possède une écriture de type phonétique (Pognan, 1999) car tous les graphèmes sont prononcés. Le français, en revanche, n'établit pas de lien direct entre le graphème et le son et pour pouvoir lire en français, il faut donc posséder des connaissances orthographiques plus complexes.

Bon nombre de chercheurs et professeurs de LE s'accordent à dire que l'initiation à une LE devrait être strictement orale car les documents écrits auraient une influence négative sur l'apprentissage phonétique. Ainsi les apprenants, ayant appris une LE avec des documents écrits, risquent de lire le graphème utilisé en LE par le son que ce symbole représente dans leur LM (Flege, 1987a; Flege, 1987d). Ainsi Delattre (1944, p. 111) précise que :

Les lettres de l'orthographe- et même les symboles d'une notation phonétique- rappellent toujours à l'étudiant les sons de sa langue maternelle ... Seul celui qui apprend directement par les sons arrive à les reproduire sans les associer avec ceux de sa langue maternelle.

Si l'on en croit Billières (1987, p. 18 - 19) :

L'écrit est particulièrement perturbateur. Des élèves qui commencent l'étude d'une langue vivante avec des méthodes faisant coïncider l'apprentissage de l'oral avec celui de l'écrit, mettent incontestablement davantage de temps à acquérir une bonne prononciation [...]. La lecture provoque des interférences qui conduisent à une sclérose rapide du système phonétique de la langue [...]. Le passage à la lecture et à l'écriture provoque un net appauvrissement de la prosodie et une tendance plus marquée à la déformation des phonèmes.

De son côté, Borrell (1996) fait la différence entre les erreurs de production induites par la mauvaise perception de la part de l'apprenant et celles qui sont causées par la graphie.

Pour leur part, Harmegnies *et al.* (2005) soulignent d'excellents résultats chez les Japonais, apprenant l'anglais, qui commencent l'apprentissage par une initiation strictement orale.

Enfin, Escudero and Wanrooij (2010) examinent l'influence de l'orthographe sur la perception des voyelles non-natives. 204 hispanophones, apprenant le néerlandais, ont passé un test de

discrimination catégorielle AXB de six voyelles du néerlandais [ɑ, a, ʏ, y, ɪ, i] dans deux conditions : sans et avec représentation graphique des voyelles perçues. Les résultats du test varient selon les conditions : dans la tâche auditive pure, il s'avère plus facile de discriminer y/i mais plus difficile de discriminer a/ɑ. Par conséquent, le fait d'ajouter le graphème correspondant au son entendu facilite la discrimination du contraste a/ɑ mais rend plus difficile à percevoir le contraste i/y. Ainsi, l'influence de l'écrit sur la perception est démontrée.

D'autres chercheurs, moins nombreux, pointent le bienfait des documents écrits utilisés dès le début de l'apprentissage. Azoulay-Vicente (1990) décrit alors les conditions d'un tel apprentissage : il faut faire systématiquement une corrélation entre les graphèmes et les sons, et expliciter les règles nécessaires pour arriver à la forme phonétique finale.

5.7 Facteurs personnels

Certains apprenants d'une langue seconde souhaitent garder leur accent étranger pour préserver leur identité (Wachs, 2011). D'autres, au contraire, veulent conserver cet accent car il peut être perçu comme noble. Il s'agit essentiellement des accents européens, notamment l'accent français, qui dénotent la sophistication et le prestige (Derwing and Munro, 2009).

La capacité à apprendre à parler comme un natif serait également liée aux traits de caractère des individus. Les études de Taylor *et al.* (1971) ont montré que l'empathie et l'intuition favorisent la production authentique. Plus on est sensible aux comportements et sentiments des autres, plus on a de chance de remarquer les subtilités qui différencient la parole étrangère de la parole maternelle.

La motivation joue également un rôle important dans l'acquisition de la parole authentique. Wachs (2011) définit trois types de motivation : celle des adultes qui veulent s'intégrer dans une nouvelle culture (intégratrice), celle des enfants qui se trouvent entourés d'enfants parlant une autre langue (assimilatrice) et celle des professionnels qui ont besoin d'acquérir la nouvelle langue pour travailler (instrumentale). Selon le type de motivation qui pousse à apprendre une nouvelle langue, on viserait une prononciation plus ou moins authentique.

5.8 Perception par le natif

La perception de l'accent étranger revient également aux auditeurs à qui les non-natifs s'adressent. Certains auditeurs seraient plus tolérants que d'autres (Munro and Derwing, 1995a), en fonction de leur propre vécu linguistique ; il en ressort qu'un ressortissant de culture plurilingue serait plus compréhensif qu'une personne monolingue (Els and Bot, 1987).

5.9 Propriétés phonético-phonologiques des sons

Les facteurs liés au locuteur et à l'auditeur ne suffisent pas souvent à expliquer le niveau de difficulté ou de facilité qu'un apprenant affronte pour acquérir certains contrastes et sons non-natifs. En effet, la difficulté réside également dans les sons eux-mêmes et il semble que certains sons soient plus faciles à acquérir que d'autres (Polka and Werker, 1994).

5.9.1 Indices acoustiques

Un autre facteur qui influence la perception des contrastes non-natifs chez les apprenants adultes est de type psycho-acoustique (Polka, 1991). En effet, les indices acoustiques, responsables du contraste perçu entre deux sons, peuvent être plus ou moins proéminents pour l'oreille humaine (Flege,

1989). Ainsi, Burnham (1986) distingue entre les contrastes robustes et fragiles. Les contrastes robustes ont une base psychoacoustique forte et ils sont perceptivement saillants. Ils sont véhiculés par exemple par les indices de voisement, de durée ou de VOT et ils sont universellement préférés (Escudero and Polka, 2003). Dans certaines langues, ces indices peuvent avoir une importance perceptive même s'ils n'ont pas de fonction véritablement phonologique. Ainsi, le système vocalique anglais oppose [i] et [ɪ] par les différences spectrales, mais se sert de l'indice de durée dans la perception (Gottfried and Beddor, 1988). Les contrastes fragiles, quant à eux, sont basés, par exemple, sur des indices spectraux, et ils sont plus difficiles à discriminer (Burnham, 1986).

5.9.2 Marque

Dans la continuité des travaux de Lado, Eckman (1977) élabore la théorie de *markedness* qui explique les écarts de l'apprenant d'une L2 en se basant sur la notion de la « marque ». Les trois postulats de cette théorie peuvent être résumés comme suit :

1. Les structures de la L2 qui sont différentes de la L1 et en même temps plus marquées, seront difficiles à acquérir.
2. Les structures de la L1 qui sont différentes mais ne sont pas plus marquées ne seront pas forcément difficiles.
3. Le niveau de difficulté effectif dépendra de la marque.

Ainsi, l'explication de l'apprentissage n'est pas uniquement d'ordre inter-langue, c'est-à-dire en comparant la L1 avec la L2, mais également d'ordre intra-langue, c'est-à-dire similaire à l'apprentissage de la LM par un enfant et la notion de difficulté s'explique par des universaux linguistiques.

La théorie de la marque appliquée à l'acquisition de la L2 permet de prédire l'évolution d'un apprenant du niveau débutant au niveau avancé (Eckman, 1985). Ce dernier acquiert d'abord les structures moins marquées, donc moins complexes avant d'acquérir les structures plus complexes. Du moment où l'apprenant réussit à acquérir la structure la plus marquée, les autres structures moins marquées pourront être maîtrisées. En revanche l'inverse n'est pas vrai : l'acquisition d'une structure moins marquée ne permet pas une généralisation de l'apprentissage aux structures plus marquées. Selon Mayerthaler (1988) les éléments marqués sont plus difficilement perceptibles et par conséquent moins bien décodés que les éléments non marqués.

5.10 Conséquences de l'accent étranger

Traditionnellement, on considère que l'accent étranger fait baisser l'intelligibilité et la compréhensibilité de la parole (Bent and Bradlow, 2003; Flege, 1988a). Derwing and Munro (2009) définissent l'intelligibilité comme le degré de compréhension. Elle peut être mesurée par des tests de retranscription de ce qui est dit. La compréhensibilité correspond à la facilité ou difficulté avec laquelle l'auditeur traite la parole de son interlocuteur. C'est le ressenti de l'auditeur qui peut être exprimé sur une échelle de jugement.

De nombreuses études (Derwing and Munro, 2009; Derwing *et al.*, 1998; Munro and Derwing, 1995a) montrent que dans des conditions parfaites, la présence de l'accent étranger jugé même très prononcé, ne diminue pas forcément l'intelligibilité et la compréhensibilité et les observations de van Wijngaarden *et al.* (2002) confortent cette idée. Els and Bot (1987) soulignent que la parole d'un locuteur non-natif, marquée par accent étranger, est en revanche sensible aux distorsions, telles que le bruit. Dans des conditions qui ne sont pas parfaites, elle est par conséquent moins intelligible que la parole native.

En revanche McAllister (1998) et (2000) montre que quoique l'accent étranger, dans des conditions idéales, ne diminue pas nécessairement l'intelligibilité auprès des auditeurs natifs, il peut

entraîner une baisse d'intelligibilité auprès des auditeurs non natifs. L'intelligibilité est d'autant plus aléatoire que la LM de l'auditeur est différente de celle du locuteur. Il s'agit d'une situation assez fréquente ; l'on est souvent amené à communiquer avec des étrangers dans une langue qui n'est ni la nôtre ni la leur et l'on a plus de chance d'être compris si l'on possède une prononciation neutre, non marquée.

Les conséquences de l'accent étranger peuvent se situer également au niveau social, psychologique ou communicatif (Derwing and Munro, 2009). Parler avec un accent peut conduire à des attitudes négatives (Flege, 1995). Un accent peut engendrer un comportement raciste, une discrimination de la part des auditeurs ou il peut être irritant (Munro and Derwing, 1995b). Oyama (1976) évoque pour sa part une forme de pénalisation sociale : le non-natif peut s'écarter de la société de peur de se faire remarquer par son accent étranger.

6. Méthodologie

Beaucoup de gens, peu d'idées, et comment faire pour nous différencier les uns des autres ?

Milan Kundera

Résumé : Le sixième chapitre explicite le protocole expérimental. Nous avons constitué trois groupes de locutrices : (1) Dix Françaises natives non-méridionales (productions en français), (2) Dix Tchèques natives de Bohême (productions en tchèque), (3) Dix futures enseignantes de FLE de LM tchèque (productions en français et en tchèque). Le corpus est composé de voyelles prononcées en isolation et dans des logatomes CVCVCVC où C est la consonne [p, t, k, ʁ] pour le français et [p, t, k, fi] pour le tchèque. Le corpus français est issu d'un corpus plus large défini par Landron *et al.* (2010). Ensuite, nous décrivons la passation des expériences en production et perception, le traitement des données acoustiques et leur visualisation. Enfin, nous détaillons les fonctionnalités du logiciel VisuVo (Visualisation des Voyelles) que nous avons conçu dans le cadre de ce travail. Il s'agit d'un système expert qui, à partir d'une grande base de données, permet de générer en instantané trois types de graphes représentant les formants vocaliques. L'innovation de cet outil par rapport aux logiciels du marché tels que Praat (Boersma and Weenink, 1992-2011), qu'il complète, réside dans l'interaction en temps réel à partir d'un site web.

6.1 Locutrices

« An attempt should be made to form subject groups that are as homogeneous as possible ... The practice of most investigators is to obtain data under identical conditions from groups of 6-12 subject. Examining fewer subjects leads to the risk of failing to observe a systematic phonetic differences between groups because of intersubject variability. »

(Flege, 1987b, p. 288)

Nous avons créé trois groupes constitués de dix locutrices féminines, âgées entre 21 et 48 ans dont la composition est illustrée à la Figure 31.

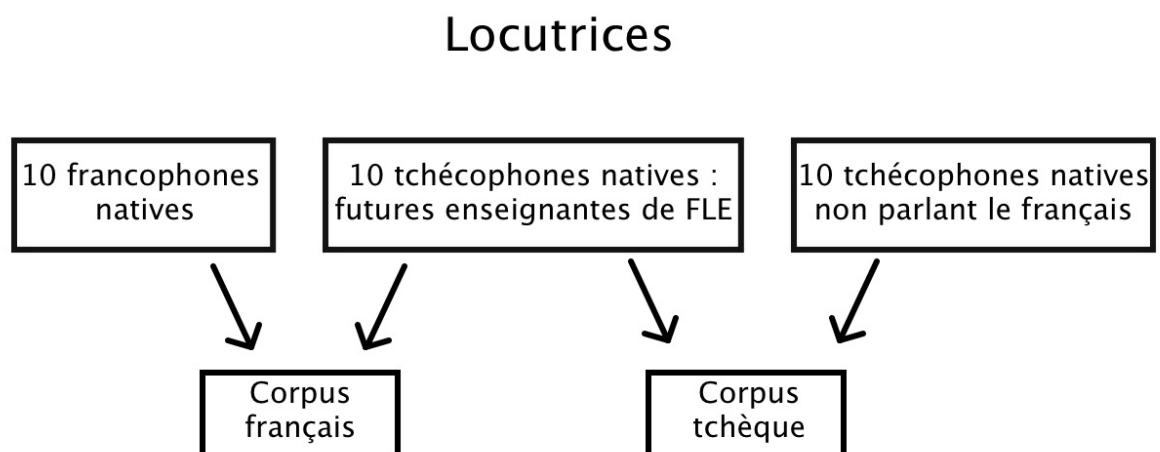


Figure 31 : Locutrices ayant produit le corpus français (10 francophones natives et 10 tchécophones natives - futures enseignantes de FLE de LM tchèque) et le corpus tchèque (10 tchécophones natives - futures enseignantes de FLE de LM tchèque et 10 tchécophones natives non parlant le français)

Le corpus français a été produit par :

1. Dix francophones natives non-méridionales, résidant au moment des enregistrements dans la région parisienne
2. Dix tchécophones natives de la région de Bohême - futures enseignantes de FLE, résidant au moment des enregistrements à Plzen (Bohême de l'ouest)

Le corpus tchèque a été produit par :

1. Dix tchécophones natives de la région de Bohême - futures enseignantes de FLE, résidant au moment des enregistrements à Plzen (Bohême de l'ouest)
2. Dix tchécophones natives de la région de Bohême, n'ayant jamais appris le français

6.1.1 Francophones natives

Les francophones natives ayant produit le corpus sur les voyelles du français sont 10 femmes, âgées entre 21 et 48 ans, venant pour la plupart de la région parisienne ou du Nord de la France ou ayant vécu au moins 6 ans à Paris. Les locutrices ont toutes fait des études supérieures dans une des universités parisiennes. Les informations les concernant se trouvent dans le Tableau 23.

Locutrice	Âge	Région où les locutrices ont grandi	Nb d'années passées en Île-de-France
F1	30	Loire Atlantique	7
F2	24	Ile-de-France	22
F3	48	Lorraine	30
F4	30	Rhône-Alpes	8
F5	30	Ile-de-France	30
F6	24	Bretagne	6
F7	26	Lorraine	6
F8	21	Ile-de-France	21
F9	26	Ile-de-France	26
F10	26	Picardie	6
Moyenne (écart type)	28,5 (7,5)		16,2 (10,5)

Tableau 23 : Informations concernant les locutrices françaises natives indiquant leur âge (avec la moyenne et l'écart type entre parenthèses), la région où elles ont grandi et le nombre d'années passées en Ile de France (avec la moyenne et l'écart type entre parenthèses)

6.1.2 Tchécophones natives, futures enseignantes de FLE

Les tchécophones natives T1 à T10 ayant enregistré le corpus tchèque et le corpus français sont des étudiantes en Pédagogie de FLE à l'Université de Bohême de l'Ouest (situé à Plzeň) et possèdent donc un niveau du français avancé. Une fois les études de Master 2 terminées avec succès, elles pourront enseigner le FLE à l'école primaire et secondaire. Comme illustré dans le Tableau 24,

- elles sont âgées entre 25 et 28 ans (la moyenne d'âge étant de 25,7 ans),
- elles ont grandi dans la région de Bohême,
- elles ont vécu entre 25 et 28 ans en région de Bohême (la moyenne étant de 25,5 ans),
- elles ont commencé à apprendre le français à l'école entre 11 et 17 ans (la moyenne d'âge étant de 14,5 ans),
- elles ont étudié le français pendant 8 à 14 ans (la moyenne étant de 10 ans),
- elles ont séjourné entre 0 et 17 mois en France (la moyenne étant de 5,6 mois),
- seulement 4 locutrices ont pour but de parler le français comme les natifs.

Les informations sur les locutrices se trouvent dans le Tableau 24 et des informations supplémentaires se trouvent dans les questionnaires qu'elles ont remplis, en annexe du chapitre 6 (pages 15 - 34 des annexes).

Locutrice	Âge	Région où les locutrices ont grandi	Nb d'années passées en Bohême	Âge au début de l'apprentissage du français	Nb d'années d'étude du français	Séjour en France (en mois)	But : prononciation native
T1	25	Bohême	25	11	14	5	non
T2	28	Bohême	28	15	10	0	non
T3	26	Bohême	26	14	10	0	non
T4	25	Bohême	25	16	8	7	oui
T5	25	Bohême	25	15	10	3	non
T6	25	Bohême	25	15	10	7	oui
T7	25	Bohême	25	15	8	0	non
T8	25	Bohême	24	17	8	12	non
T9	25	Bohême	25	13	12	5	oui
T10	28	Bohême	27	14	10	17	oui
Moyenne (écart type)	25,7 (1,3)		25,5 (1,2)	14,5 (1,6)	10 (1,9)	5,6	

Tableau 24 : Informations concernant les futures enseignantes de FLE de LM tchèque indiquant l'âge au début de l'apprentissage du français, le nombre d'années d'étude du français (avec les moyennes et les écarts types entre parenthèses) et le nombre de mois passés en France

6.1.3 Tchécophones natives, non parlant le français

Un autre groupe de tchécoslovaques natives, n'ayant jamais étudié le français, ont enregistré le corpus sur les voyelles du tchèque. Il s'agit de 10 femmes, âgées de 24 à 38 ans (la moyenne étant de 28,3 ans). Elles sont nées et ont vécu dans la région de Bohême et elles ont effectué des études supérieures dans une des universités de Bohême. Ces informations se trouvent dans le Tableau 25.

Locutrice	Age	Région où les locutrices ont grandi	Nb d'années passées en Bohême
T11	27	Bohême	27
T12	28	Bohême	27
T13	27	Bohême	26
T14	28	Bohême	26
T15	28	Bohême	28
T16	38	Bohême	38
T17	27	Bohême	27
T18	28	Bohême	28
T19	24	Bohême	24
T20	28	Bohême	28
Moyenne (écart type)	28,3 (3,6)		27,9 (3,8)

Tableau 25 : Informations concernant les locutrices tchèques natives indiquant leur âge (avec la moyenne et l'écart type entre parenthèses), la région où elles ont grandi et le nombre d'années passées en Bohême (avec la moyenne et l'écart type entre parenthèses)

6.2 Corpus

Nous avons utilisé une partie du corpus PhoDiFLE (Landron *et al.*, 2010), construit par « le Groupe Didactique » du LPP de Paris 3 que nous avons fondé. Ce corpus permet une comparaison de données des chercheurs qui travaillent sur des problématiques liées à l'enseignement de la prononciation du FLE. Les parties du corpus utilisées dans cette thèse portent sur les voyelles orales en isolation et en contexte. Pour étudier les voyelles monophthongues du tchèque, le corpus est adapté de celui utilisé pour le français.

6.2.1 Voyelles isolées

Dans le corpus sur le français, dix voyelles orales [i, e, ε, y, ø, œ, u, o, ɔ, a] sont insérées dans des phrases cadre du type « Papa, il a dit <a> comme dans papa. ». Le fait de rajouter le mot « papa » en début de phrase permet de préparer le locuteur à la voyelle qu'il devra prononcer isolément. Nous sommes consciente que les voyelles isolées n'apparaissent que très rarement dans la parole spontanée (Meunier, 2001) et que les voyelles moyennes se trouvent en distribution semi-complémentaire

(Wioland, 2005). Or, le choix du corpus sur les voyelles isolées est néanmoins pertinent car il permet de définir acoustiquement la cible vocalique nécessaire pour l'étude de l'influence du contexte consonantique sur la réalisation acoustique de la voyelle (Meunier, 2001) et l'analyse acoustique montre que les locuteurs natifs sont capables de produire les dix différentes qualités vocaliques en isolation, tout comme les Français natifs de l'étude de Gottfried (1984).

Dans le corpus tchèque, les dix voyelles monophthongues [ɪ, i:, ɛ, ɛ:, u, u:, o, o:, a, a:] sont insérées dans des phrases cadre du type « *Jana, říkám <a> jako ve slově Jana.*¹⁰ ».

Toutes les phrases sont mélangées et sont répétées quatre fois durant la séance d'enregistrement.

6.2.2 Voyelles en contexte dans trois positions du logatome

Le corpus combine les dix voyelles orales du français [i, e, ɛ, y, ø, œ, u, o, ɔ, a] avec les consonnes de différents lieux d'articulation (labial, dental, palato-vélaire, uvulaire) [p, t, k, ɣ], dans trois syllabes différentes (initiale, médiane, finale), créant ainsi des logatomes trisyllabiques qui sont insérés dans des phrases cadre du type « Le mot papapape peut bien coller. »

<p style="text-align: center;"><u>Corpus français</u></p> <p style="text-align: center;">CViCVmCVfC</p> <p style="text-align: center;">où C = [p, t, k, ɣ]</p> <p style="text-align: center;">V = [i, e, ɛ, y, ø, œ, u, o, ɔ, a]</p> <p style="text-align: center;">i = initiale, m = médiane, f = finale</p>
--

À côté de chaque phrase, nous avons ajouté entre crochets le symbole phonétique de la voyelle cible, à l'intention de ceux qui maîtrisent ces symboles, et des dessins de mots qui contiennent cette voyelle (par exemple le dessin d'une pelle pour la voyelle [ɛ]). Voir la Figure 32 (page 111) pour illustration.

Dans le corpus tchèque, nous combinons les dix voyelles monophthongues [ɪ, i:, ɛ, ɛ:, u, u:, o, o:, a, a:] avec les consonnes de différents lieux d'articulation (labial, dental, vélaire, glottal) [p, t, k, h] (le phonème /ɣ/ n'existant pas dans le système phonologique du tchèque, il est remplacé par la consonne glottale /h/). Les logatomes ainsi créés sont également trisyllabiques, du type CVCVCVC, et ils apparaissent dans des phrases cadre telles que « *Slovo papapap pŭsobí divně*¹¹. ».

<p style="text-align: center;"><u>Corpus tchèque</u></p> <p style="text-align: center;">CViCVmCVfC</p> <p style="text-align: center;">où C = [p, t, k, h]</p> <p style="text-align: center;">V = [ɪ, i:, ɛ, ɛ:, u, u:, o, o:, a, a:]</p> <p style="text-align: center;">i = initiale, m = médiane, f = finale</p>
--

¹⁰ Nous traduisons « Jana, je dis <a> comme dans le mot Jana. »

¹¹ Nous traduisons « Le mot papapap donne une impression bizarre. »

Toutes les phrases sont mélangées et sont répétées quatre fois durant la séance d'enregistrement.

6.3 Passation de l'expérience en production

6.3.1 Phase d'entraînement

Avant de commencer l'enregistrement, les locutrices prennent connaissance du corpus et s'entraînent. Ainsi, toutes les locutrices, qu'elles soient natives ou non, passent par une phase d'entraînement qui est constituée d'un visionnage d'une courte présentation dans PowerPoint où certaines consignes sont écrites, d'autres enregistrées dans un fichier .wav. Aucune consigne n'est donnée oralement par l'expérimentateur surveillant les enregistrements afin que toutes les locutrices obtiennent exactement les mêmes informations. Des exemples sonores des phrases qui apparaissent dans le corpus sont fournis. Les consignes générales, présentées avant l'enregistrement des deux corpora (sur les voyelles isolées et les voyelles en contexte), sont les suivantes :

- | |
|---|
| <ul style="list-style-type: none">- Lire ni trop lentement ni trop rapidement.- Lire avec un débit normal.- Lire chaque phrase de façon continue, sans marquer de pauses.- Lire chaque phrase d'abord dans sa tête avant de la lire à haute voix pour éviter une lecture trop hâtive.- Lire calmement. On ne vous évalue pas, votre participation est très précieuse pour nous. |
|---|

Les consignes spécifiques au corpus sur les voyelles isolées sont les suivantes :






















- | |
|--|
| <ul style="list-style-type: none">- Les phrases qui vous seront proposées sont construites de la manière suivante :- Exemple : Ville, il a dit <i> comme dans ville- La voyelle qui apparaît entre crochets est celle du mot de référence qui apparaît au début et à la fin de la phrase. Détachez bien cette voyelle du mot précédent et suivant- Essayez de lire cette phrase et on vous dira si vous avez bien compris le fonctionnement |
|--|

Les consignes spécifiques au corpus sur les voyelles en contexte sont les suivantes :

- | |
|--|
| <ul style="list-style-type: none">- Les phrases qui vous seront proposées contiennent des mots inexistants- Exemple des mots inexistants : dédedéde, goeugoeugoeugue- Pour lire ces mots inventés, appliquez les règles de lecture telles que vous les avez apprises pour des mots existants- Dans les mots inventés, ne lisez pas le dernier <e>. Exemple : le mot « dédedéde » doit être lu [dededed] en non pas [dedededø] |
|--|

Des exemples de diapositives de la phase d'entraînement des voyelles en contexte phonétique se trouvent à la Figure 32.

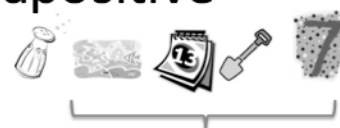
Comment dois je lire la voyelle ? A quel son correspond-t-elle ?

Voyelle (graphème)	Mot inexistant avec la voyelle	Symbole phonétique de la voyelle	Mots existants avec cette voyelle	Dessins correspondant aux mots
<i>	<u>bibibibe</u> 	[i]	Lit, six, mille 	  1000
<é>	<u>tétététe</u> 	[e]	Thé, blé, fée 	  
<è>	<u>mèèmème</u> 	[ɛ]	Sel, mer, pelle, sept, treize 	    
<a>	<u>papapape</u> 	[a]	Rat, chat, quatre 	  

Exemple d'une diapositive

[ɛ]

Symbole phonétique de la voyelle <è>



Dessins des mots existants qui contiennent la voyelle <è>

Voyelle <è>

Le mot sèssèssesse peut bien coller.

Mot inexistant

Figure 32 : Phase d'entraînement des voyelles du français insérées en contexte phonétique

6.3.2 Phase d'enregistrement

L'enregistrement du corpus commence immédiatement après l'entraînement et il se déroule dans une pièce calme. Les phrases à lire sont présentées dans un document PowerPoint où chaque diapositive ne contient qu'une seule phrase. Les locutrices avancent d'une diapositive à l'autre à leur rythme, en appuyant sur une flèche du clavier de l'ordinateur. Après chaque série de phrases (au total

quatre séries car quatre répétitions), une diapositive propose aux locutrices de marquer une pause. Un verre d'eau est à leur disposition tout au long de l'enregistrement.

Le matériel d'enregistrement est composé d'un microphone serre-tête AKG C520 et d'une carte-son Edirol UA25. Les phrases sont enregistrées sur l'ordinateur portable Mac via le logiciel AudaCity 2.0.4 (<http://audacity.sourceforge.net/>). La fréquence d'échantillonnage est fixée à 44100 Hz et la quantification est de 16 bits.

6.4 Base de données

Les fichiers sons sont stockés au format WAV (extension .wav). Nous extrayons les parties du signal correspondant aux voyelles en isolation et aux logatomes CVCVCVC pour les soumettre à l'analyse acoustique. Notons qu'un certain nombre de données aberrantes dues à une mauvaise détection de formants est écarté.

6.4.1 Le français par dix Françaises natives

La base de données sur le français est constituée de 399 voyelles isolées. Sur les 400 voyelles isolées à enregistrer (10 voyelles orales * 4 répétitions * 10 locutrices), nous avons écarté une occurrence du [e] et deux occurrences du [i]. Nous avons en revanche enregistré et ajouté une occurrence de [ø] et une occurrence de [u] supplémentaires (l'une des locutrices a produit ces deux voyelles cinq fois au lieu de quatre fois).

La base de donnée sur le français contient ensuite 4793 voyelles en contexte. Sur les 4800 voyelles en contexte à enregistrer (10 voyelles orales * 4 contextes * 3 positions dans le logatome * 4 répétitions * 10 locutrices), nous avons écarté :

- deux logatomes [papapap], soit 6 voyelles
- un logatome [ɛiɛiɛi], soit 3 voyelles
- un logatome [ɔøɔøɔ], soit 3 voyelles
- deux logatomes [kækækæk], soit 6 voyelles
- un logatome [kykyky], soit 3 voyelles
- une voyelle finale du logatome [ɛiɛiɛi]

A l'inverse, nous avons enregistré des répétitions supplémentaires de :

- deux logatomes [kækækæk], soit 6 voyelles
- deux logatomes [køkøkøk], soit 6 voyelles
- un logatome [kækækæk], soit 3 voyelles

Il est à noter que nous n'avons pas pris en considération la valeur finale des trois formants d'une occurrence du [e] de la syllabe initiale du logatome [tetetet], la valeur finale des deux formants d'une occurrence du [u] de la syllabe finale du logatome [pupupup] et enfin la valeur finale des trois formants d'une occurrence du [y] de la syllabe finale du logatome [tytytyt].

6.4.2 Le tchèque par 20 Tchèques natives

La base de données du tchèque contient 800 voyelles isolées (10 voyelles orales * 4 répétitions * 20 locutrices) et 9600 voyelles en contexte (10 voyelles orales * 4 contextes * trois positions * 4 répétitions * 20 locutrices). Il est à noter que nous avons écarté une fois la valeur finale des deux formants d'une occurrence du [o:] de la syllabe finale du logatome [to:to:to:t], la valeur finale des deux formants d'une occurrence du [u] de la syllabe finale du logatome [ɦuɦuɦuɦ] et enfin la valeur finale des deux formants d'une occurrence du [u:] de la syllabe finale du logatome [tu:tu:tu:t].

6.4.3 Le français par 10 futures enseignantes de FLE de LM tchèque

La base de données du français constituée à partir de productions de 10 futures enseignantes de FLE de LM tchèque contient 398 voyelles isolées. Sur les 400 voyelles isolées à enregistrer (dix voyelles orales * quatre répétitions * dix locutrices), nous avons écarté une occurrence du [e] et une occurrence du [i]. La base de données contient ensuite 5081 voyelles en contexte. Sur les 4800 voyelles en contexte à enregistrer (10 voyelles orales * 4 contextes * 3 positions * 4 répétitions * 10 locutrices), nous avons écarté :

- deux logatomes [kakakak], soit 6 voyelles
- une voyelle finale du logatome [kakak**ak**]
- un logatome [tatatat], soit 3 voyelles
- deux voyelles finales du logatome [ɤɤɤɤ**ɤɤ**]
- un logatome [ɤɤɤɤɤ], soit 3 voyelles
- deux logatomes [pupupup], soit 6 voyelles
- deux logatomes [tututut], soit 6 voyelles

En revanche, nous avons enregistré et analysé des répétitions supplémentaires par des locutrices T5 et T10 de :

- quatre logatomes [kekekek], soit 12 voyelles
- quatre logatomes [pepepep], soit 12 voyelles
- huit logatomes [kɛkɛkɛk], soit 24 voyelles
- huit logatomes [pɛpɛpɛp], soit 24 voyelles
- quatre voyelles initiales et médianes du logatome [ɤ**ɤɤɤɤɤ**], soit 8 voyelles
- huit logatomes [kokokok], soit 24 voyelles
- huit logatomes [popopop], soit 24 voyelles
- deux logatomes [ɤɤɤɤɤ], soit 6 voyelles
- trois logatomes [tototot], soit 9 voyelles
- sept logatomes [kɔkɔkɔk], soit 21 voyelles
- huit logatomes [pɔpɔpɔp], soit 24 voyelles
- trois logatomes [tɔtɔtɔt], soit 9 voyelles
- deux logatomes [ɤæɤæɤæɤ], soit 6 voyelles
- cinq logatomes [kækækæk], soit 15 voyelles
- sept logatomes [pæpæpæp], soit 21 voyelles
- un logatome [tætætæt], soit 3 voyelles
- huit logatomes [kɔkɔkɔk], soit 24 voyelles
- huit logatomes [pɔpɔpɔp], soit 24 voyelles
- deux logatomes [ɤɔɤɔɤɔ], soit 6 voyelles
- quatre logatomes [tɔtɔtɔt], soit 12 voyelles

Il est à noter que nous avons écarté de notre analyse :

- la valeur médiane des trois formants d'une occurrence du [a] de la syllabe finale du logatome [ɤ**ɤɤɤɤɤ**]
- la valeur finale des quatre formants d'une occurrence du [i] de la syllabe finale du logatome [k**ikikik**]
- la valeur initiale des quatre formants d'une occurrence du [i] de la syllabe finale du logatome [ɤ**ikikik**]
- la valeur initiale et une fois les valeurs médiane et finale des deux formants de deux occurrences du [u] de la syllabe finale du logatome [k**ukukuk**]

- la valeur finale des trois formants d'une occurrence de [y] de la syllabe initiale du logatome [tytytyt]

6.5 Traitement des données

Les fichiers sons contenant des voyelles isolées et des logatomes trisyllabiques sont traités dans le logiciel Praat 5.3.42 (Boersma and Weenink, 1992-2011) où nous effectuons leur segmentation et l'étiquetage de manière semi-automatique à l'aide du plugin EasyAlign (Goldman, 2011). La position des frontières est vérifiée et en cas de besoin ajustée manuellement. Pour la segmentation des voyelles, nous nous appuyons sur l'oscillogramme et le spectrogramme et la partie choisie est celle où les quatre premiers formants (notamment le F2) sont clairement visibles.

Ensuite, les valeurs des formants sont relevées à environ un tiers, à la moitié et deux tiers de la durée vocalique en utilisant le script semi-automatique Log_File_4 de C. Gendrot du LPP de Paris 3 (<http://gendrot.ilpga.fr/scripts.htm>). Ces trois moments de prise de valeurs sont importants notamment pour l'étude de la coarticulation dont les effets sont visibles non seulement durant les transitions formantiques mais également dans la partie stable de la voyelle (Durand, 1985; Hillenbrand *et al.*, 2001). La Figure 33 illustre ce phénomène.

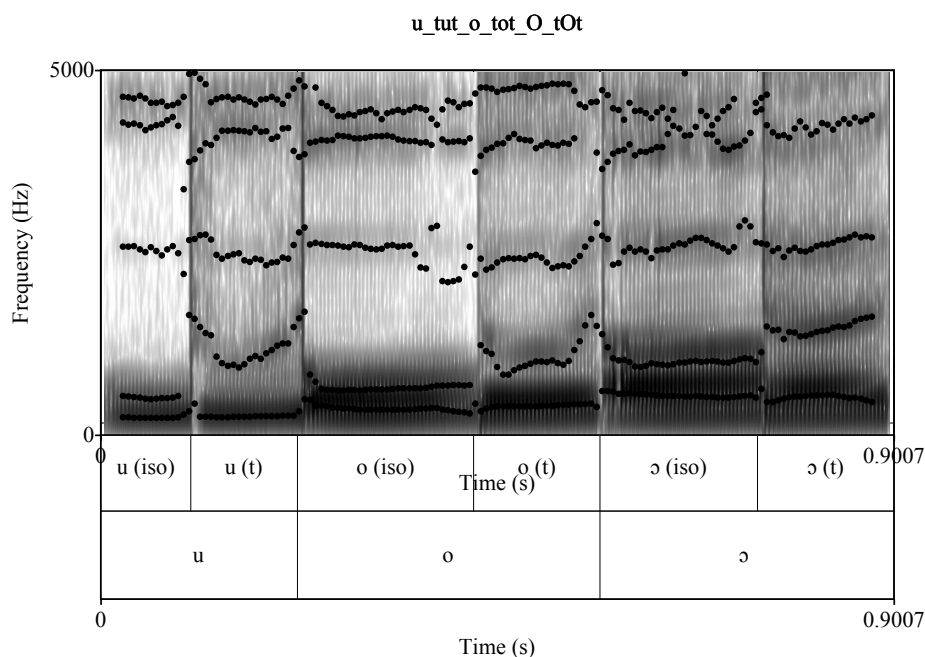


Figure 33 : Spectrogramme des voyelles françaises [u, o, ɔ], prononcées par une Française native en isolation (iso) et en contexte symétrique dental (t)

L'avantage de l'application d'un script semi-automatique est qu'il permet de relever la valeur du formant à l'endroit où aucune anomalie du mouvement formantique n'est constatée (anomalies qui pourraient provenir par exemple de périodes de glottalisation ou *voice decay time*). La formule utilisée « Get formant » repose sur l'analyse LPC¹² qui fixe la largeur de la fenêtre à 25 ms afin de pouvoir visualiser les formants. La détection des cinq premiers formants se fait entre 0 et 5500 Hz (et non 5000 Hz, fréquence couramment employée par défaut pour la recherche des formants dans les voix

¹² L'analyse LPC (Linear Predictive Coding) est basée sur le modèle source-filtre. Elle fournit de bons résultats pour les voyelles orales mais n'est pas adaptée à l'analyse des voyelles nasales (se caractérisant par des anti-formants).

masculines car les fréquences de résonance des locutrices sont plus élevées que celles des hommes). Les valeurs sont directement stockées dans un fichier .xls qui sert d'entrée au logiciel VisuVo (Visualisation des Voyelles) permettant la visualisation des résultats. Ce programme est décrit plus en détail dans la section 6.6 qui suit.

Comme rappelé en 1.3, bon nombre de publications, notamment des chercheurs travaillant sur l'anglais, se cantonnent aux valeurs des deux premiers formants pour comparer toutes les voyelles, comme si les deux premiers formants étaient suffisants pour représenter le timbre de toutes les voyelles, ce qui est une trop grande simplification (cours de Master de Phonétique de J. Vaissière).

Le timbre des voyelles orales est fidèlement reproduit par la resynthèse des quatre premiers formants. Pourtant tous les formants ne jouent pas un rôle identique pour l'identification des voyelles : les deux premiers formants suffisent pour certaines voyelles, mais pas pour toutes (Delattre *et al.*, 1951).

Dans cette thèse, nous avons comparé les voyelles postérieures par leurs deux premiers formants, les voyelles antérieures par les trois premiers formants et enfin le [i] par les quatre premiers formants (le rapprochement de F3/F4 étant nécessaire pour créer une concentration d'énergie vers 3200 Hz). Nous avons montré que ces formants sont suffisants pour la description acoustique dans un test d'identification de cinq séries de dix voyelles isolées du français prononcées par cinq Françaises natives non-méridionales où seuls les formants jugés a priori nécessaires ont été conservés. La description et les résultats du test sont détaillés dans les annexes (pages 38 -42 de l'annexe du chapitre 6).

6.6 Visualisation des résultats acoustiques

Le logiciel VisuVo, que nous avons conçu dans le cadre de cette thèse, est un outil de visualisation des formants des voyelles.

Techniquement, le logiciel VisuVo est composé de deux parties majeures : la partie Serveur et la partie Client. La partie Serveur permet de pré-calculer les données fournies en entrée par un fichier Excel (tel que nous l'avons obtenu à partir de Praat par exemple)¹³ et elle donne la moyenne et l'écart type à partir de valeurs de chaque onglet séparément. Les moyennes peuvent alors concerner plusieurs groupes de locuteurs préalablement déterminés (comme les Français, les Tchèques ou bien toutes les futures enseignantes) ou bien être dégagées à partir de productions d'un seul locuteur. Il revient à l'utilisateur de décider avec quelles données il veut travailler et de répartir les valeurs dans des onglets en fonction de l'objectif de son travail.

Une fois les données analysées et converties en un arbre d'objets par le serveur, la partie Client peut générer les graphes en temps réel. Cette partie est écrite en langage Javascript et s'appuie sur le moteur Javascript disponible sur le navigateur (Google Chrome ou Safari par exemple). Le Client tire ensuite partie des fonctionnalités HTML5 de rendu Canvas disponibles sur les navigateurs modernes pour générer des graphes de haute qualité qui sont de trois types : (1) Le premier permet d'afficher et de comparer les différentes voyelles d'une ou de plusieurs langues et il correspond au triangle vocalique classique (F1/F2 et F2/F3). (2) Le deuxième montre l'évolution des formants d'une ou de plusieurs voyelles suivant la position prosodique et l'entourage consonantique (Figure 34 à la page 116). (3) Le troisième met en évidence les contrastes entre les voyelles isolées d'une langue (Figure 35 à la page 117). A la différence d'un logiciel tel que Praat qui est certes puissant mais très lent pour générer une image à partir d'une grande base de données, VisuVo permet à l'utilisateur d'afficher les

¹³ Notons que le serveur VisuVo peut analyser une très grande quantité de données en moins de deux secondes. A titre d'information, le serveur a été utilisé pour analyser un fichier Excel de 12 onglets comprenant 81754 lignes sur 16 colonnes de données, soit environ 1,3 million de données.

données de façon interactive et immédiate. Comme illustré à la Figure 34, l'utilisateur manipule dans les trois types de graphiques les paramètres suivants :

1. Le locuteur ou groupe(s) de locuteurs dont il souhaite étudier les productions. Le choix dans notre thèse est le suivant : tchèque (groupe de 20 natives), français (groupe de 10 natives), locutrices T1, T2, T3 ... T10 (futures enseignantes de FLE individuellement), 1-10 (groupe de 10 futures enseignantes),
2. l'écart type affiché,
3. l'unité (Hertz, Bark, ERB),
4. le mode d'affichage : mode noir et blanc ou couleur,
5. la prise d'images, stockées ensuite dans le serveur : oui ou non,
6. l'affichage de la valeur moyenne et de l'écart type exact lorsqu'il clique sur un symbole vocalique précis,
7. l'exportation dans un fichier Excel séparé des valeurs formantiques moyennes calculées et affichées par le logiciel.

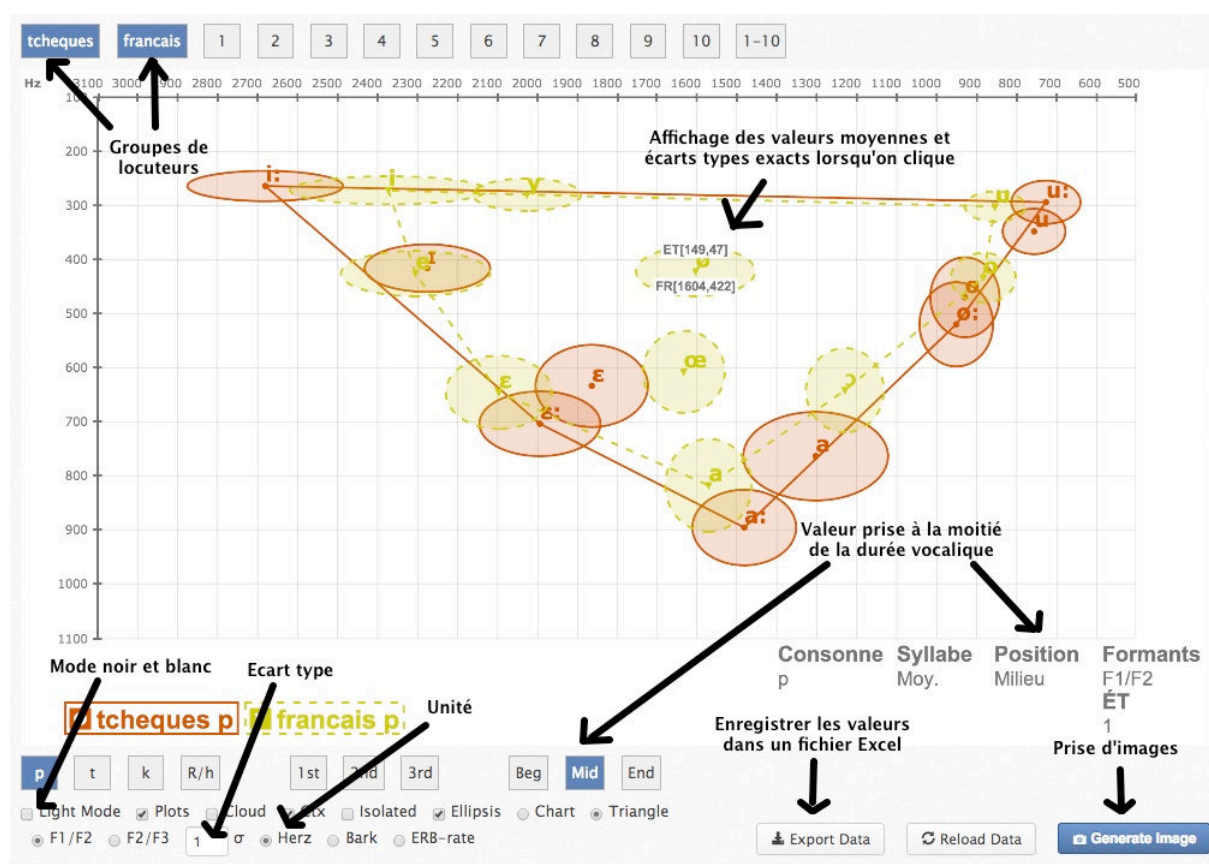


Figure 34 : Fonctionnalités générales de la partie Client du logiciel VisuVo

Outre les fonctionnalités décrites ci-dessus, l'utilisateur détermine dans le graphique du triangle vocalique les facteurs suivants :

- Les axes (F1/F2 ou F2/F3),
- l'affichage des voyelles en isolation (bouton « Isolated ») et/ou produites en contexte (« Ctx »)
- si « Ctx » coché, l'utilisateur choisit l'entourage consonantique p, t, k, R ou h. Si aucun entourage coché, les valeurs moyennes affichées sont celles de tous les contextes mélangés,
- la position prosodique : voyelle initiale (« 1st »), médiane (« 2nd »), finale (« 3rd »). Si aucune position cochée, les valeurs moyennes affichées sont celles de toutes les positions,
- le moment de la prise de valeur : début (« Beg »), milieu (« Mid ») ou fin (« End ») vocalique. Si aucun moment de la prise de valeur coché, les valeurs moyennes affichées sont calculées à partir des trois valeurs prises par voyelle,

- l’affichage du nuage des points (« Cloud ») : pour visualiser la position acoustique exacte de tous les sons produits qui ont servi à calculer la moyenne.

Pour passer aux deuxième et troisième types de graphique, l'utilisateur appuie sur le bouton « Chart ». Comme illustré à la Figure 35, l'utilisateur peut ensuite choisir la voyelle dont il veut suivre l'évolution des formants en fonction de la position prosodique et de l'entourage consonantique. Il peut également comparer les formants de différentes voyelles de différents groupes de locuteurs (ici le [i] des Tchèques avec le [e] des Français).



Figure 35 : Fonctionnalités propres au graphique « Chart » de la partie Client du logiciel VisuVo

Enfin, pour visualiser les contrastes entre les voyelles isolées d'une langue, l'utilisateur désactive dans le graphique « Chart » l'affichage des voyelles en contextes « Ctx » et ne garde que l'affichage des voyelles isolées (en cochant « Isolated »). Il peut modifier l'ordre de leur présentation en modifiant l'ordre des boutons du bas de même que zoomer pour ne visualiser que le formant F1 par exemple. Ces fonctionnalités sont illustrées à la Figure 36 à la page 118.

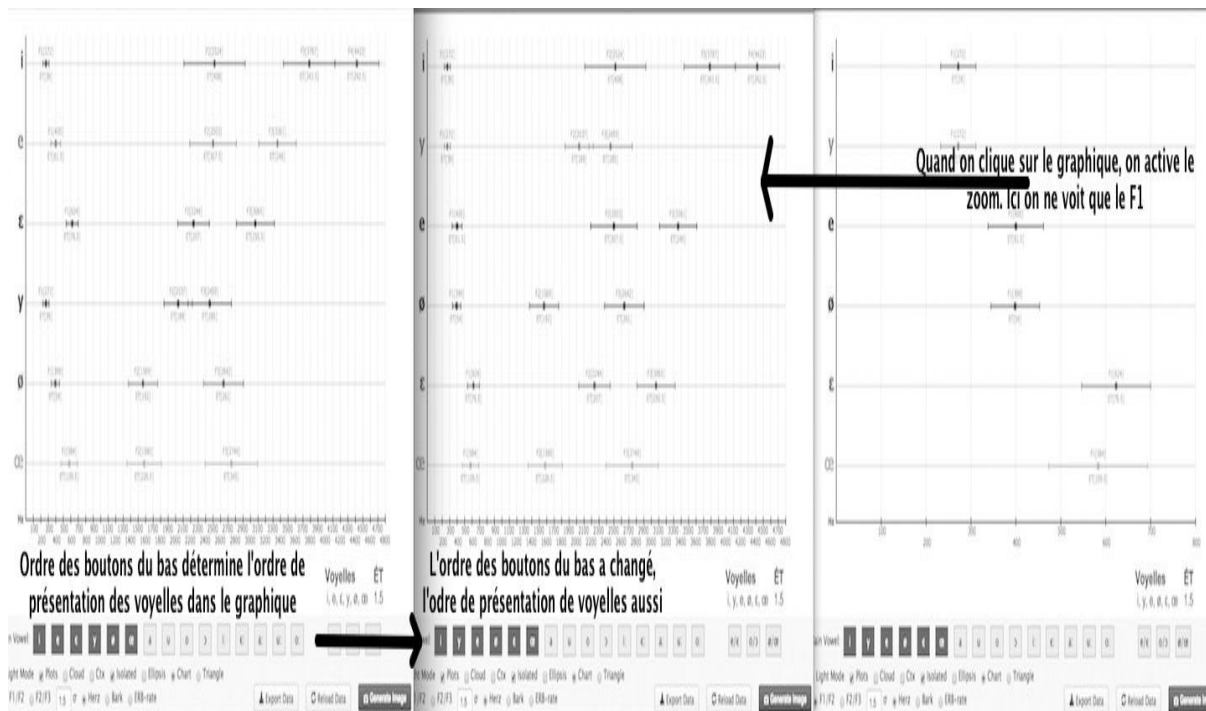


Figure 36 : Fonctionnalités propres au graphique « Chart » des voyelles isolées de la partie Client du logiciel VisuVo

Même s'il est possible d'installer VisuVo sur un poste de travail comme une application « classique », l'avantage principal de VisuVo est d'être installé sur un serveur, potentiellement partagé par plusieurs utilisateurs et de se lancer depuis un navigateur web comme n'importe quel site web. Son utilisation ne requiert donc aucune installation pour l'utilisateur final.

Le calcul de l'écart type, les formules de conversion et la méthode de dessin sont plus largement décrits en annexe du chapitre 6, pages 37 - 38.

6.7 Interprétation des résultats acoustiques

Afin de quantifier la similarité acoustique entre les voyelles des deux langues et pour indiquer le degré d'authenticité des voyelles produites par les futures enseignantes de FLE de LM tchèque, nous mesurons la distance des formants vocaliques en termes d'écart types. Les voyelles postérieures sont comparées avec les deux premiers formants, les voyelles antérieures avec les trois premiers formants et les [i, i:, ɪ] avec les quatre premiers formants. Le critère pour définir la similarité et l'authenticité acoustique est représenté dans le Tableau 26.

Critère acoustique	Similarité acoustique entre les voyelles du français et du tchèque	Niveau d'authenticité acoustique des voyelles du français par des tchécophones par rapport à la « référence » (= 10 francophones natives)
Pas de différence significative des formants comparés selon le test-t ($p > 0,05$)	Voyelle <i>identique</i>	Voyelle à priori authentique
Distance des formants comparés inférieure ou égal à 1 écart type	Voyelle « hautement similaire » Similarité +++	Voyelle à priori authentique ¹⁴
Distance des formants entre 1 et 2 écarts types : - Distance des formants entre 1 – 1,5 écarts types - Distance des formants entre 1,5 – 2 écarts types	Voyelle <i>similaire</i> - Similarité ++ - Similarité +	Voyelle à priori non authentique
Distance d'au moins un formant comparé supérieure à 2 écarts types	Voyelle <i>nouvelle</i> Similarité 0	Voyelle à priori non maîtrisée

Tableau 26 : Critères acoustiques utilisés dans notre thèse pour définir les voyelles *identiques*, « hautement similaires », *similaires* et *nouvelles* de même que les voyelles authentiques/ non authentiques, non maîtrisées

¹⁴ L'authenticité doit être confirmée par des tests de perception

Le critère qui repose sur la valeur d'un écart type dans la définition des sons « hautement similaires »/ authentiques est celui de Birdsong (2003) qui mesure l'authenticité de la prononciation du français des apprenants anglophones tardifs par rapport à la « référence » ou celui de Guion *et al.* (2000) qui indiquent la similarité perceptive entre les sons de l'anglais et du japonais.

Ainsi, afin de quantifier le degré de similarité acoustique interlangue, les règles sont les suivantes :

1. Les voyelles du français dont tous les formants comparés sont identiques à ceux des voyelles tchèques (sans différence significative des formants, avec $p > 0,05$) sont appelées acoustiquement *identiques*.
2. Puisque les voyelles *identiques* entre le français et le tchèque n'existent pas, comme nous allons le voir en 8.4 et 9.5, nous introduisons le terme « hautement similaires » (avec « similarité +++ »), déjà utilisé par Bohn et Flege (1990). Nous caractérisons ces voyelles par des formants se trouvant à un écart type ou moins de la moyenne et donc dans une plage de normalité avec un niveau de confiance de 68%.
3. Les voyelles françaises dont les formants se trouvent entre un et deux écarts types de ceux des voyelles tchèques les plus proches acoustiquement sont alors plus ou moins *similaires*. « La similarité ++ » caractérise alors les voyelles dont les formants se trouvent entre 1 – 1,5 écarts types et la « similarité + » décrit les voyelles dont les formants se situent entre 1,5 – 2 écarts types.
4. Enfin les voyelles françaises dont un ou plusieurs formants se trouvent à plus de deux écarts types de ceux des dix voyelles tchèques sont considérées comme acoustiquement *nouvelles* (« similarité 0 »).

De la même manière, afin de quantifier le niveau de maîtrise des voyelles du français produites par les futures enseignantes de FLE de LM tchèque, les règles sont les suivantes :

1. Les voyelles produites avec tous les formants comparés à moins d'un écart type de ceux de la « référence » sont considérées comme acoustiquement *authentiques*.
2. Les voyelles produites avec au moins l'un des formants comparés entre un et deux écarts types de ceux de la « référence » ne sont pas authentiques. Nous considérons cependant que la catégorie vocalique est maîtrisée et que la difficulté se trouve au niveau phonétique.
3. Les voyelles produites avec au moins l'un des formants à plus de deux écarts types de la « référence » ne sont pas maîtrisées, la difficulté peut se trouver au niveau phonologique car l'écart par rapport à la référence peut induire un changement de sens.

Notons néanmoins que l'authenticité doit être confirmée par des tests de perception effectués par des auditeurs français natifs.

6.8 Tests de perception et leur interprétation

Des différences acoustiques entre deux groupes que des tests statistiques indiquent comme significatives peuvent cependant ne pas être perceptibles ou perçues.

Nous avons ainsi effectué plusieurs séries d'épreuves de perception :

1. Test d'identification de voyelles du français avant et après suppression de formants supérieurs à F2 pour les voyelles postérieures, à F3 pour les voyelles antérieures et à F4 pour [i] (afin de vérifier que les formants maintenus suffisent pour la définition acoustique des voyelles en question),
2. test d'identification de voyelles longues et brèves du tchèque raccourcies à 150 ms de la partie centrale (pour décider si la voyelle brève et son homologue longue de chaque couple ont le même timbre vocalique),

6. Méthodologie

3. test d'identification interlangue des voyelles orales du français par des auditeurs tchèques monolingues (dans le but de tester la similarité perceptive entre les voyelles du français et du tchèque),
4. test d'identification des voyelles françaises produites par des Françaises natives et identifiées par les futures enseignantes de FLE de LM tchèque (afin de vérifier l'existence de la catégorie perceptive pour les dix voyelles du français chez les tchécophones),
5. test d'identification des voyelles françaises produites par les futures enseignantes de FLE de LM tchèque et identifiées par des auditeurs français natifs (afin de vérifier la maîtrise/l'authenticité des voyelles produites par les non-natives),
6. une série de tests de perception des voyelles moyennes produites et perçues par des Français natifs (afin de vérifier la sensibilité des natifs aux contrastes entre les voyelles moyennes) dont un test d'identification de sons vocaliques synthétisés avec Klatt,
7. test d'identification des voyelles moyennes du français produites par les futures enseignantes de FLE de LM tchèque (10 répétitions) et identifiées par des auditeurs français natifs (afin de vérifier la maîtrise/l'authenticité des voyelles).

Bien que les stimuli et les auditeurs varient d'un test à l'autre, la procédure de construction et les conditions de passation des tests restent à peu près identiques : l'auditeur est placé dans une pièce calme et identifie les stimuli perçus. L'écoute se fait avec un casque ; la plupart des auditeurs utilisent le modèle Quite Bose que nous mettons à leur disposition.

Les tests 1 à 5 sont construits avec le logiciel Praat, qui permet d'enregistrer automatiquement dans des fichiers sortant la réponse de l'auditeur et la note de qualité qu'il attribue à chaque stimulus. L'échelle de justesse s'étend dans tous les tests de 1 à 5 où 5 correspond à la meilleure note attribuée. Dans les tests 2 et 3, les notes correspondent à :

1. voyelle très mal prononcée
2. voyelle mal prononcée
3. entre les deux
4. voyelle bien prononcée
5. voyelle très bien prononcée

Dans les tests 1, 4, 5, 6 et 7, les notes expriment le niveau de confiance, à savoir:

1. j'en suis sûr(e)
2. plutôt sûr(e)
3. entre les deux
4. plutôt pas sûr(e)
5. je n'en suis pas sûr(e)

Les consignes varient d'un test à l'autre et elles sont précisées dans la partie expérimentale. Les tests commencent par des stimuli d'entraînement qui permettent à l'auditeur de comprendre ce qui est attendu de lui et peuvent inclure des stimuli distracteurs. Dans les tests 1 à 5, chaque voyelle est répétée deux fois et les deux répétitions sont séparées par un silence d'une seconde. Ainsi, en cas de manque de concentration, la voyelle est automatiquement rejouée ce qui nous a permis d'omettre la possibilité de réécouter le stimulus à volonté. Des pauses sont régulièrement insérées tous les 15 stimuli. Le nombre de stimuli varie d'un test à l'autre mais le nombre maximal est de 139.

Les résultats des tests d'identification sont retranscrits dans des tableaux où la première colonne renseigne sur les stimuli, la deuxième sur la (les) réponse(s). La troisième colonne indique ensuite le pourcentage d'identification et la quatrième la note de confiance ou de jugement. La dernière colonne fournit « l'index de correspondance » (*le fit index*) introduit dans les travaux de Guion *et al.* (2000, p. 2716). Cette valeur permet, selon les auteurs, de combiner l'identification et l'évaluation de la justesse en une seule métrique.

Le *fit index* s'étend alors de 0 à 5. Ainsi plus la valeur du *fit index* baisse, plus la similarité interlangue, la maîtrise ou la capacité à identifier la voyelle baisse.

L'avantage de recourir à des tests d'identification dans les études qui s'intéressent à l'acquisition phonétique d'une L2 se trouve dans :

- la vérification de la perceptibilité des différences acoustiques (et donc de « l'accent étranger ») trouvées dans les données de la production (lorsque les tests sont effectués par les auditeurs natifs),
- la possibilité d'étudier chez les apprenants de FLE le rapport de cause à effet entre perception et productions (lorsque les tests sont effectués par des non-natifs).

6.9 Le choix du modèle d'apprentissage phonétique

Pour prédire et expliquer l'apprentissage phonétique des voyelles du français par les futures enseignantes de LM tchèque, nous avons adopté le Speech Learning Model (SLM) de Flege (1995). Ce modèle est particulièrement adapté à notre étude car :

- il vise des apprenants tardifs de niveau avancé
- il s'intéresse non seulement à la perception des apprenants mais également à leur production
- il n'explique pas tous les écarts en production par une perception erronée de l'apprenant

Le choix de ce modèle par rapport aux autres modèles actuels, dont les hypothèses sont exposées en 3.4, est discuté en conclusion finale de cette thèse.

Deuxième partie : Comparaison des voyelles orales du français non-méridional et du tchèque de Bohème

« To accurately gauge the L2 learner's success in modifying previously established patterns of production (or perception) of phones in the L2, it is first necessary to document differences between L1 and L2 phones. »

(Flege, 1987b, p. 288)

Dans la deuxième partie (chapitres 7 - 11), nous comparons les voyelles françaises et tchèques de trois points de vue :

- a) les symboles phonétiques utilisés pour leur transcription,
- b) les propriétés physiques (formants et durée) et
- c) la similarité perceptive (par des tests de perception interlangue).

7. Symboles vocaliques (API)

The IPA was originally designed for transcribing phonemic oppositions. The articulatory characterization that it provides is not precise enough to pinpoint a specific vowel quality... the same symbol occasionally receives contradictory characterizations.

(Vaissière, 2011, p. 53)

Résumé : Le **chapitre sept** compare les symboles phonétiques utilisés pour décrire les systèmes vocaliques des deux langues. Les symboles de l'API que nous avons choisis pour décrire le système phonétique du tchèque sont les suivants : [a, a:, ɪ, i, u, u:, ε, ε: o, o:]. Ce choix est en concordance avec la littérature récente (Dankovicova, 1997a). Cependant, en considérant également les autres symboles utilisés pour décrire le système vocalique tchèque, seules les voyelles du français [y, ø, œ] seraient *nouvelles* alors que [a, i, u, e, ε, o, ɔ] seraient *similaires* ou *identiques* aux voyelles du tchèque.

L'API a été conçu à l'origine pour faciliter l'apprentissage des langues, comme exposé en 4.1. Dans le cas de la transcription des voyelles, chaque symbole est une abstraction du timbre concret de la voyelle et il représente les multiples allophones d'un même phonème. Les mêmes symboles, comme souvent /a/, /e/, /i/, /o/, /u/, sont utilisés pour décrire les systèmes phonologiques des langues à cinq voyelles, sans prendre en compte la réalisation exacte du son dans chaque langue (Vaissière, 2006). Dans ce sens, l'utilisation de l'API peut induire une certaine ambiguïté et complexifier l'apprentissage du timbre exact en L2 plutôt que le simplifier (Vaissière, 2007).

Alors qu'il existe un consensus sur le choix des symboles phonétiques de l'API pour transcrire les voyelles du français, la transcription phonétique de certaines voyelles du tchèque varie d'un auteur à l'autre. Les deux triangles vocaliques représentant les voyelles du français et du tchèque sont illustrés à la Figure 37. Les symboles alternatifs utilisés pour transcrire une même voyelle tchèque sont indiqués entre parenthèses.

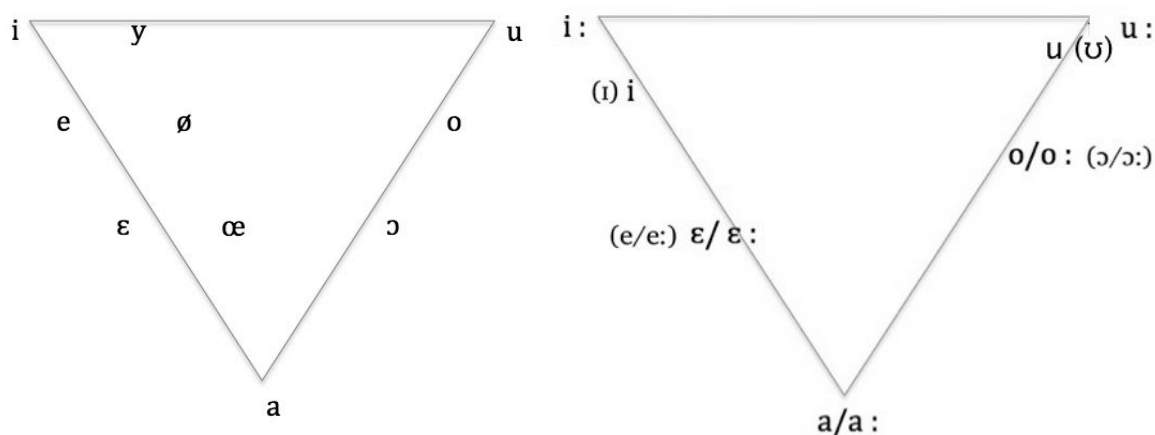


Figure 37 : Triangle vocalique du français standard actuel (à gauche) et du tchèque de Bohême (à droite), les symboles divergents sont entre parenthèses

Les axes d'un triangle vocalique représentent à peu près les valeurs attendues du premier formant, lié principalement au trait d'aperture et deuxième formant, lié principalement au trait d'antériorité-postériorité. Notons que la labialité joue un rôle important sur les premier et second formants. Les voyelles nasales du français /ã, ê, ï/ et les voyelles diphtongues du tchèque /au, eu, ou/ ne sont pas représentées à la Figure 37. Cependant les voyelles nasales du français sont souvent

confondues avec des voyelles orales même si elles sont différentes des points de vue articulaire et acoustique (Delvaux *et al.*, 2002). Ces dernières posent également des problèmes pour leur apprentissage.

7.1 En français

Le système phonologique des voyelles du français est composé, selon les régions et les pays, de 13 à 15 voyelles (Vaissière, 2006). Le système vocalique du français standard actuel compte treize voyelles : dix voyelles orales /i, e, ε, a, u, o, ɔ, y, ø, œ/ et trois voyelles nasales /ĩ, ê, õ/. Cependant l'opposition entre les voyelles moyennes e/ε, o/ɔ et ø/œ est neutralisée en position inaccentuée et il convient pour des raisons didactiques de transcrire les voyelles /e, ε/ en syllabe non-finale du mot phonétique par l'archiphonème /E/, /o, ɔ/ par /O/ et /ø, œ/ par /œ/ (Wioland, 2005).

Le système vocalique maximal comprend en plus le /ɑ/ postérieur opposant « Anne » et « âne » et /œ̃/ opposant « brun » et « brin » qui est un contraste encore réalisé dans les parlers du sud de la France (Léon and Léon, 2007).

Enfin, quoique le schwa /ə/ soit inclus par de nombreux auteurs dans le système phonologique du français, son caractère phonémique est discutable. Cette voyelle apparaît en effet seulement en position inaccentuée et elle peut disparaître (Wioland, 2001).

Ainsi les voyelles monophthongues du tchèque seront comparées avec dix voyelles orales du français retranscrites par les symboles phonétiques [i, e, ε, a, u, o, ɔ, y, ø, œ].

7.2 En tchèque

- L'inventaire vocalique du tchèque est traditionnellement composé de dix voyelles monophthongues /i, i:, e, e:, o, o:, u, u:, a, a:/ avec cinq « catégories spectrales »¹⁵, phonologiquement brèves et longues (Kucera and George, 1968; Ludvikova and Kraus, 1966; Mazlova, 1946).
- Le timbre des trois voyelles extrêmes longues du tchèque de Bohême est toujours noté par les symboles phonétiques [a:, i:, u:].
- L'équivalent bref de [a:] est retranscrit avec [a] par tous les auteurs.
- Les homologues brèves des voyelles [i:] et [u:] sont en revanche retranscrits respectivement par [i] ou [ɪ] puis [u] ou [ʊ] pour montrer en plus de la différence de quantité également la différence de qualité (Dankovicova, 1997a; Dubeda, 2005; Dubeda and Januska, 2006; Simackova *et al.*, 2012; Skarnitzl and Volin, 2012).
- La transcription des voyelles moyennes du tchèque varie également selon les auteurs. Dans la description du tchèque ayant un système vocalique à trois degrés d'aperture, les chercheurs doivent choisir parmi les symboles offerts par l'API et donc d'indiquer soit la qualité mi-fermée, soit la qualité mi-ouverte. Traditionnellement, Hála (1941) ou Borovičková and Maláč (1967) utilisent les symboles correspondant aux voyelles mi-fermées [e] (et [e:]) ainsi que [o] (et [o:]). Plus récemment, les auteurs retranscrivent la voyelle antérieure moyenne avec le symbole mi-ouvert [ɛ] pour la voyelle brève et [ɛ:] pour la voyelle longue (Dankovicova, 1997a; Dubeda, 2005; Dubeda and Januska, 2006; Simackova *et al.*, 2012; Skarnitzl and Volin, 2012). De même, la voyelle postérieure moyenne est retranscrite par certains auteurs avec le symbole mi-ouvert [ɔ] (Dubeda, 2005; Dubeda and Januska, 2006; Podlipsky, 2009).

¹⁵ Terme „spectral categories“ utilisé par Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A. and Nishi, K. (2001). Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *Journal of the Acoustical Society of America*, 109.

Notre propre choix des symboles de transcription phonétique des voyelles tchèques est en concordance avec la littérature récente (Dankovicova, 1997a) et correspond aux [a:, u:, i:, ε:, o:] pour les voyelles longues et aux [a, u, i, ε, o] pour les voyelles brèves.

Les symboles [ε/ε:] marquant la qualité mi-ouverte, [o/o:] marquant la qualité mi-fermée et [ɪ] marquant le caractère relâché sont choisis suite aux résultats de nos analyses acoustiques et perceptives des voyelles tchèques, exposés dans les chapitres 8 et 9 ci-dessous et sont en accord avec la littérature récente.

Ainsi, en se basant uniquement sur le critère des symboles phonétiques utilisés par les différents auteurs, les voyelles du français *similaires* ou *identiques* aux voyelles tchèques sont [a, i, u, e, ε, o, ɔ] et les voyelles *nouvelles* [y, ø, œ].

8. Propriétés acoustiques des voyelles isolées

Nous considérons la production de voyelles isolées comme un contexte particulier, plutôt rare, mais dont les caractéristiques ciblées présentent l'avantage d'être une sorte de degré zéro de la variation contextuelle.

(Meunier, 2001, p. 78)

Résumé : Dans le **chapitre huit**, nous étudions le patron formantique des voyelles isolées du français et du tchèque qui représentent, selon Miller (1981), la cible acoustique. Les valeurs cibles servent de références pour l'étude des déviations formantiques dues au phénomène de coarticulation des voyelles avec les consonnes. Cette définition des valeurs dites cibles est à prendre toute fois avec précaution, du moins pour le français. En effet les voyelles moyennes mi-ouvertes /ɔ, œ/ ne sont jamais distribuées en syllabe finale ouverte en français standard et la production de ces voyelles en isolation enfreint donc les règles orthoépiques (Léon, 1976). Il a également été noté par Vaissière (dans ses cours de phonétique expérimentale de Master) que les Français natifs prononcent en général un [ɑ] pharyngal (F1 et F2 proches) en isolation et un [a] central (F2 autour de 1500 Hz) en contexte.

Ensuite, nous mesurons les distances acoustiques entre les formants successifs des voyelles dites focales. Nous étudions également les indices acoustiques des contrastes vocaliques d'aperture, d'antériorité-postériorité, de labialité et de durée.

La comparaison des formants vocaliques montre qu'il n'existe pas de voyelles acoustiquement *identiques* entre le français et le tchèque, quoique le [ɑ] français soit par ses formants *hautement similaire* au [a] tchèque (les deux ayant un F2 inférieur à 1500 Hz et étant focales en isolation). Les voyelles du français [y, ø, œ] sont ensuite acoustiquement *nouvelles* car elles sont réalisées dans une zone F1/F2 inoccupée par les voyelles du tchèque. Les voyelles restantes [i, u, e, ε, o, ɔ] sont *similaires* aux voyelles tchèques.

8.1 Valeurs formantiques des cibles

Les valeurs formantiques moyennes des voyelles cibles du français et du tchèque ont été calculées sur les stimuli suivants :

- quatre répétitions de chacune des dix voyelles orales du français [i, e, ε, y, ø, œ, u, o, ɔ, a] produites par dix Françaises natives non-méridionales
- quatre répétitions de chacune des dix voyelles monophthongues du tchèque [ɪ, i:, ε, ε:, u, u:, o, o:, a, a:] produites par vingt Tchèques natives de la région de Bohême

Les valeurs utilisées pour le calcul des formants moyens des voyelles isolées ont été relevées à environ un tiers, à la moitié et à deux tiers de la durée vocalique. Le corpus utilisé et le calcul des valeurs moyennes sont décrits dans le chapitre 6 sur la méthodologie et ils sont récapitulés dans le Tableau 27 de la page 130.

CORPUS	
français	10 voyelles orales du français [i, e, ε, y, ø, œ, u, o, ɔ, a], insérées dans des phrases cadre. exemple : « Papa, il a dit <a> comme dans papa. », 4 répétitions
tchèque	10 voyelles monophthongues du tchèque [ɪ, i:, ε, ε:, u, u:, o, o:, a, a:], insérées dans des phrases cadre exemple : « <u>Jana</u> , říkám <a> jako ve slově Jana. ¹⁶ », 4 répétitions
LOCUTRICES	
françaises	10 Françaises non-méridionales natives ayant grandi ou longtemps séjourné en Ile-de-France
tchèques	20 Tchèques natives de la région de Bohême
MESURES	
Calcul de la moyenne formantique	A partir de 3 valeurs, relevées à un tiers, à la moitié, et à 2 tiers de la durée vocalique (10 ou 20 locutrices* 4 répétitions)

Tableau 27 : Corpus, locutrices pour les mesures des moyennes formantiques des voyelles isolées du français et du tchèque

Le Tableau 28 montre les valeurs moyennes en Hertz des F1 et F2 de toutes les voyelles, du F3 des voyelles antérieures, du F4 de [i, ɪ, i:] et les écarts types (exprimés en Hertz et en pourcentage). Les valeurs extrêmes sont en gras. Pour une meilleure lisibilité, les voyelles des deux langues retranscrites avec les mêmes symboles phonétiques se trouvent sur la même ligne.

¹⁶ Nous traduisons « Hana, je dis <a> comme dans le mot Hana. »

8. Propriétés acoustiques des voyelles isolées

FR	F1	F2	F3	F4	TCH	F1	F2	F3	F4
i	273 (25, 9,2%)	2524 (271, 10,7%)	3787 (227, 6%)	4428 (195, 4,4%)	i:	260 (37, 14,2%)	2734 (153, 5,6%)	3599 (185, 5,1%)	4264 (217, 5,1%)
					ɪ	403 (52, 12,9%)	2545 (133, 5,2%)	3087 (138, 4,5%)	4184 (227, 5,4%)
y	273 (25, 9,2%)	2037 (125, 6,1%)	2460 (190, 7,7%)						
e	400 (40, 10%)	2504 (205, 8,2%)	3364 (165, 4,9%)						
ø	400 (35, 8,6%)	1571 (127, 8,1%)	2645 (173, 6,5%)						
ɛ	624 (50, 8,1%)	2244 (136, 6,1%)	3066 (165, 5,4%)		ɛ	708 (45, 6,4%)	2040 (145, 7,1%)	2913 (143, 4,9%)	
					ɛ:	701 (60, 8,6%)	2094 (140, 6,7%)	2936 (155, 5,3%)	
œ	585 (72, 12,3%)	1579 (154, 9,8%)	2751 (229, 8,3%)						
ɑ	805 (130, 16,1%)	1301 (178, 13,7%)	2832 (167, 5,9%)		ɑ	869 (76, 8,7%)	1344 (101, 7,5%)	2726 (164, 6%)	
					a:	930 (82, 8,8%)	1452 (120, 8,3%)	2739 (152, 5,5%)	
u	288 (28, 9,2%)	770 (70, 9,1%)			u	341 (45, 13,2%)	718 (80, 11,1%)		
					u:	292 (39, 13,4%)	660 (79, 12%)		
o	417 (36, 8,3%)	791 (80, 10,1%)			o	540 (65, 12%)	945 (94, 9,9%)		
					o:	512 (64, 12,5%)	926 (80, 8,6%)		
ɔ	577 (80, 13,9%)	1063 (89, 8,4%)							

Tableau 28 : Valeurs moyennes (en Hz) de F1 et F2 (toutes les voyelles), F3 (voyelles antérieures) et F4 (voyelles [i, ɪ, i:]) des voyelles du français et du tchèque (3 valeurs*10 locutrices*4 répétitions). Les écarts types se trouvent entre parenthèses. Les valeurs extrêmes sont en gras. FR = français, TCH = tchèque

Le Tableau 28 montre qu'en français, les voyelles fermées et mi-fermées sont généralement et relativement moins dispersées au niveau du premier formant (avec un écart type égal ou inférieure à 10 %) que les voyelles mi-ouvertes et le [ɑ]. L'exception est la voyelle [ɛ] qui présente l'écart type de F1 et F2 le plus réduit de toutes les voyelles produites. En tchèque, la dispersion autour de la moyenne

de F1 est plus grande qu'en français ce qui s'explique sans doute par un nombre de locuteurs deux fois plus élevé. Contrairement au français, l'écart type de F1 est le moins élevé pour les voyelles moyennes et pour [a] que pour les voyelles fermées. Le [ɛ], de même qu'en français, présente la plus petite dispersion de réalisations de la valeur de F1.

La Figure 38 superpose sur l'axe F1 et F2 (en Bark) les triangles acoustiques des voyelles du français (en pointillé) et du tchèque (en trait plein) prononcées en isolation. L'écart type affiché est de un. On observe une très grande dispersion des valeurs des formants du [a] français prononcé en isolation (avec une réalisation pharyngale).

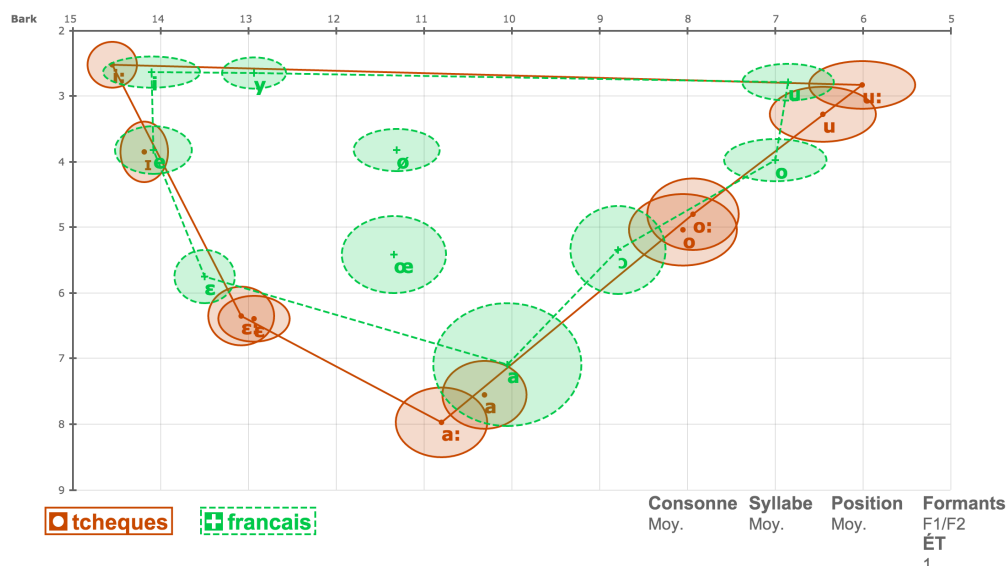


Figure 38 : Triangle vocalique des voyelles orales isolées du français (en pointillé) et du tchèque (en trait plein) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 10 Françaises/ 20 Tchèques*4 répétitions*3 valeurs par voyelle

Le F1 des voyelles longues [i:] et [a:] et le F2 de [i:] et [u:] du tchèque prononcées isolément atteignent des valeurs en moyenne proches mais souvent plus extrêmes que les voyelles extrêmes du français, [i], [u] et [a]. Ce résultat est relativement inattendu car les voyelles du français sont souvent considérées comme des voyelles particulièrement extrêmes. Le F1 de [i, a, u] est respectivement de 273 Hz, 805 Hz, 288 Hz pour le français et de 260 Hz (donc - 13 Hz), 930 Hz (+ 125 Hz) et 292 Hz (+ 4 Hz) pour [i:, a:, u:] du tchèque. Le F2 de [u] français est de 770 Hz et celui de [u] tchèque est de 718 Hz (- 52 Hz). Le F2 est encore plus bas pour le [u:] long tchèque qui est de 660 Hz (- 110 Hz). Le résultat montrant un F2 plus bas pour le [u] tchèque par rapport au [u] français est en accord avec Dohalska-Zichova (1980). Le F2 de [i] est de 2524 Hz pour le français et celui de [i] tchèque est de 2734 Hz (+ 210 Hz). En revanche, F3 est le plus élevé pour [i] français : il est de 3787 Hz en moyenne alors que celui de [i:] tchèque est de 3599 Hz (- 188 Hz).

Les valeurs extrêmes de certains formants des voyelles orales isolées du français [i, y, u, a] sont visibles à la Figure 39 qui est un spectrogramme de [i, e, ɛ, a, ɔ, o, u, œ, ø, y] prononcées par une Française native parisienne : [i] avec F3 maximal, [a] avec F1 maximal, [u] avec F2 minimal, et [y] avec F3 minimal. Les valeurs formantiques de ces voyelles sont détaillées et commentées en annexe du chapitre 8, page 43 des annexes.

8. Propriétés acoustiques des voyelles isolées

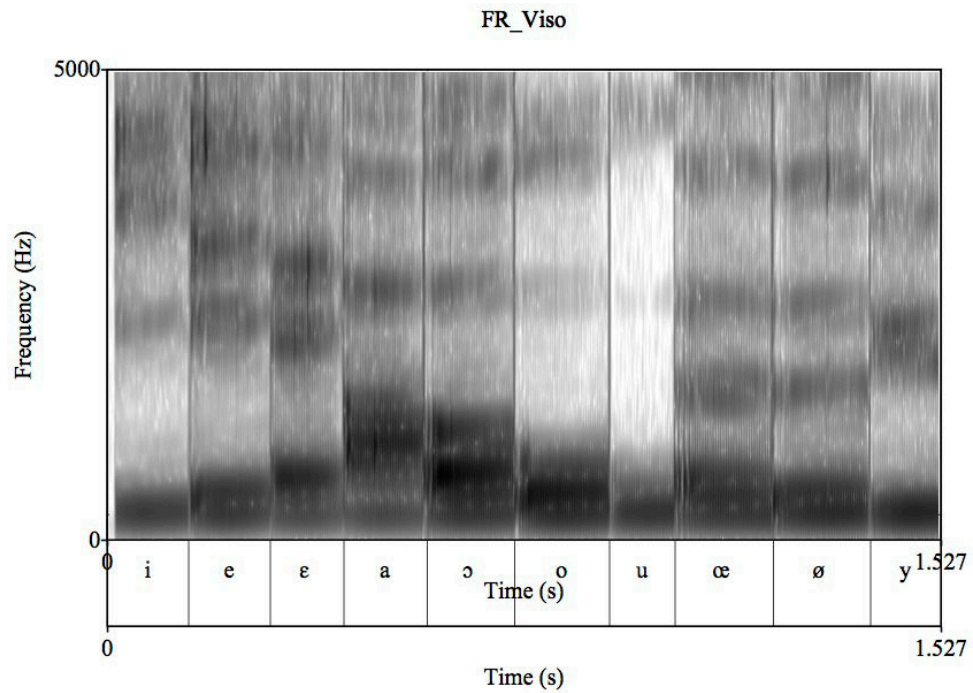


Figure 39 : Spectrogramme des voyelles orales isolées [i, e, ε, a, ɔ, o, u, œ, ø, y] prononcées par une Française native

Dans le cas des voyelles tchèques, les valeurs formantiques minimales et maximales de [i:, a:, u:, ɔ:] sont illustrées à la Figure 40 qui est un spectrogramme des dix voyelles monophthongues [i:, ɪ, ε, ε:, ɔ, a:, ɪ, o:, u, u:] prononcées par une Tchèque native. Il est à noter à la Figure 40 que les voyelles brèves et leurs homologues longues possèdent un timbre proche, à l'exception des voyelles de la paire i/i: où le [ɪ] bref se caractérise par un F1 plus élevé et un F2 et F3 plus bas que le [i:] long. Les valeurs formantiques extrêmes sont listées en annexe du chapitre 8, page 43 des annexes.

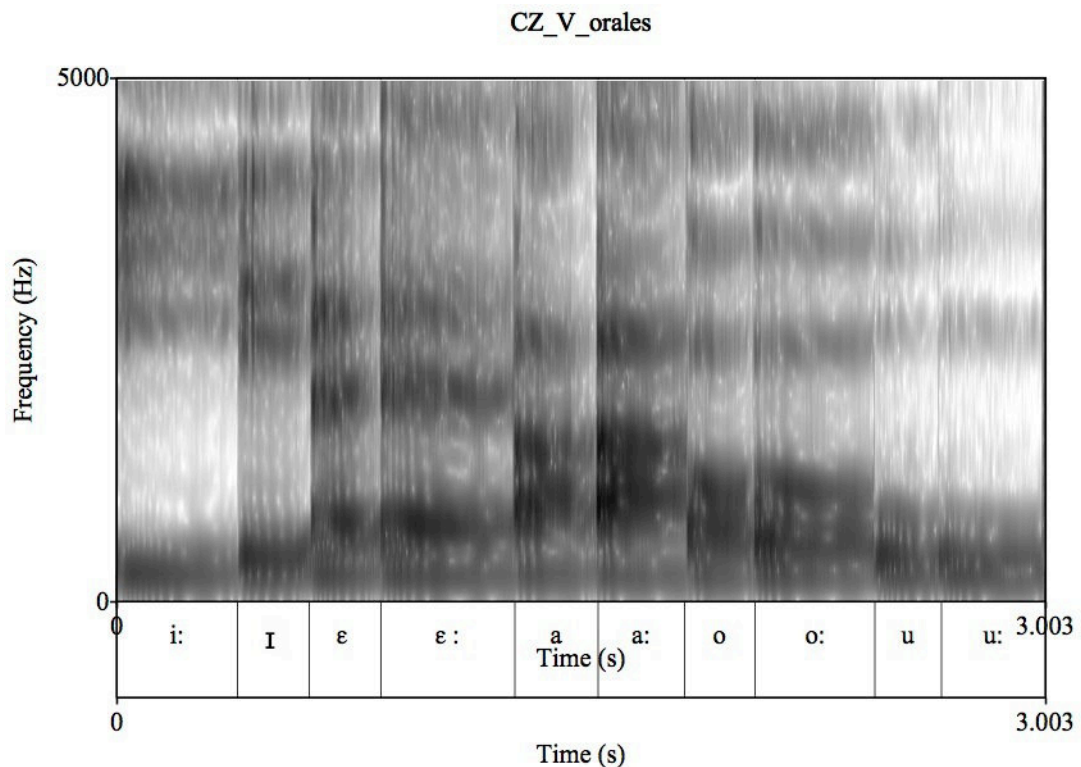


Figure 40 : Spectrogramme des voyelles tchèques [i:, ɪ, ε, ε:, ɔ, a:, ɪ, o:, u, u:] prononcées par une Tchèque native

Le spectrogramme de la Figure 41 met en contraste les voyelles extrêmes du français [i, a, u], prononcées par une Française native, et celles du tchèque [i:, a:, u:], prononcées par une Tchèque native. Nous y remarquons que les [i:] et [a:] longs tchèques ont un F2 plus élevé que les [i] et [a] du français et le [u:] long tchèque se réalise au contraire avec un F2 plus bas que le [u] français.

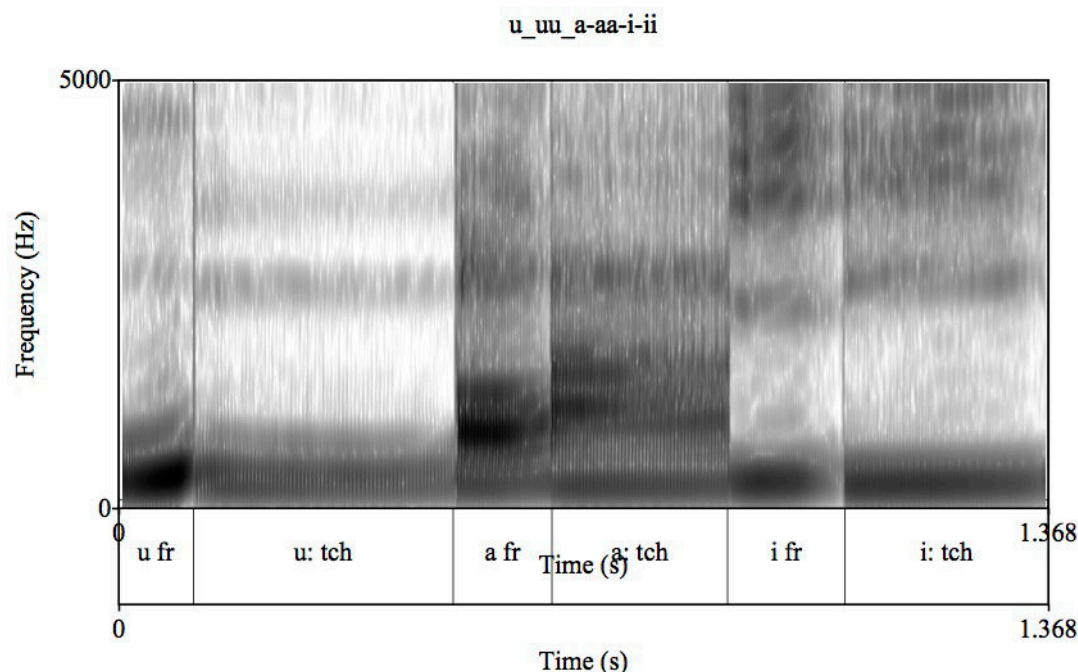


Figure 41 : Spectrogramme des voyelles [u, a, i] du français et [u:, a:, i:] du tchèque, prononcées isolément par une Française et une Tchèque native

8.2 Focalisation formantique

La focalisation des voyelles se traduit par un rapprochement de deux formants successifs, comme exposé dans le chapitre 1, page 10, qui sont amplifiés. La convergence de F1/F2 se produit dans la région près de la glotte et caractérise les voyelles postérieures alors que la convergence des formants supérieurs se produit en avant du conduit vocal et caractérise les voyelles antérieures.

Nous avons étudié la distance entre les formants successifs liée au caractère focal des voyelles du français et du tchèque en calculant la distance acoustique entre les formants rapprochés et donc amplifiés. Les distances acoustiques F2-F1 des voyelles postérieures, de F2-F1 et F3-F2 des voyelles antérieures et F2-F1, F3-F2, F4-F3 (en Hertz) du [i] français et [i:] tchèque sont indiquées dans le Tableau 29.

Le Tableau 29 montre qu'en isolation, les voyelles postérieures du français [u, o, ɔ] et du tchèque [u, u:, ɔ, ɔ:] ainsi que le [a, a:] (avec une réalisation pharyngale en isolation) sont acoustiquement focales car les formants F2-F1 sont rapprochés. La distance F2-F1 du [u] français est de 482 Hz en moyenne, alors que celle des [u, u:] tchèque est respectivement de 377 Hz et 368 Hz. La distance F2-F1 du [a] français est de 496 Hz et celle des [a, a:] tchèque est respectivement de 475 Hz et 522 Hz. Les voyelles moyennes postérieures du français et du tchèque présentent une distance acoustique F2-F1 de 374 Hz et 486 Hz pour [o, ɔ] français respectivement et de 405 Hz et 414 Hz pour [o, o:] tchèques respectivement. En ce qui concerne les voyelles antérieures arrondies, présentes dans le système français, le caractère focal de la voyelle [y] (noté par exemple par Gendrot *et al.* (2008) qui indiquent la distance F3-F2 de 622 Hz pour [y] en parole radiophonique), est confirmé. La distance acoustique F3-F2 de [y] est selon nos données de 423 Hz. En tchèque, un rapprochement de F3/F2 est remarqué pour la voyelle [ɪ] avec une distance acoustique de 542 Hz entre les deux formants. Alors qu'en français, le rapprochement de F3/F2 pour la voyelle [y] est dû à la valeur très basse du F3 (marque de labialité dans le cas des voyelles antérieures) qui est de 2460 Hz, celui de la voyelle [ɪ]

8. Propriétés acoustiques des voyelles isolées

tchèque est dû à la valeur très élevée du F2 qui est de 2545 Hz. Enfin, les voyelles focales [i] du français et [i:] du tchèque présentent une distance acoustique F4-F3 semblable, à savoir de 641 Hz et de 665 Hz respectivement. Ce rapprochement des formants F3 et F4 se produit en français dans des plus hautes fréquences qu'en tchèque : la valeur du F3 est en moyenne de 3787 Hz contre 3599 Hz pour le tchèque et celle du F4 est en moyenne de 4428 Hz contre 4264 Hz pour le tchèque.

FR	F2-F1	F3-F2	F4-F3	TCH	F2-F1	F3-F2	F4-F3
i	2251	1263	641	i:	2474	865	665
				ɪ	2142	542	1097
y	1764	423					
e	2104	860					
ø	1171	1074					
ɛ	1620	822		ɛ	1332	873	
				ɛ:	1393	842	
œ	994	1172					
a	496	1531		ɑ	475	1382	
				a:	522	1287	
u	482			u	377		
				u:	368		
o	374			o	405		
				o:	414		
ɔ	486						

Tableau 29 : Distances acoustiques F2-F1 (de toutes les voyelles), F3-F2 (voyelles antérieures et [a]) et F4-F3 (voyelles [i, ɪ, i:] des voyelles du français (FR) et du tchèque (TCH) en Hertz. Les rapprochements formantiques sont grisés et le rapprochement maximal est en gras. Les valeurs moyennes sont calculées à partir de productions de 10 locutrices*4 répétitions pour le français et 20 locutrices*4 répétitions pour le tchèque

Pour résumer, les distances acoustiques entre les formants F2-F1 des voyelles postérieures et du /ɑ/ (réalisé pharyngale en isolation) du français et du tchèque et des formants F3-F2 de [y] français sont en général inférieures à 500 Hz. En revanche, les distances F2-F1 du [a:] tchèque et F3-F2 du [ɪ] tchèque dépassent 500 Hz : il ne s'agit donc a priori pas de voyelles focales.

La distance acoustique des formants F4-F3 des voyelles isolées [i] français et [i:] tchèque se situe autour de 650 Hz. Pour comparaison, la distance F4-F3 du [i] français reportée par Gendrot *et al.* (2008) avec des mesures prises sur des voyelles dans de la parole continue est de 646 Hz. De cette façon, nous pouvons qualifier les voyelles du français [u, o, ɔ, ɑ, y, i] et du tchèque [u, u: o, o:, ɑ, i:] comme « focales ».

La représentation spectrographique des voyelles isolées du français [u, o, ɔ, a, y, i] prononcées par une Française native se trouve à la Figure 42 et celle des voyelles isolées du tchèque [u, u:, o, o:, a, a:, ɪ, i:] prononcées par une Tchèque native se trouve à la Figure 43. Notons qu'en tchèque, la distance F2-F1 du [a:] est supérieure à celle du [a] bref et que [ɪ] n'est pas une voyelle focale.

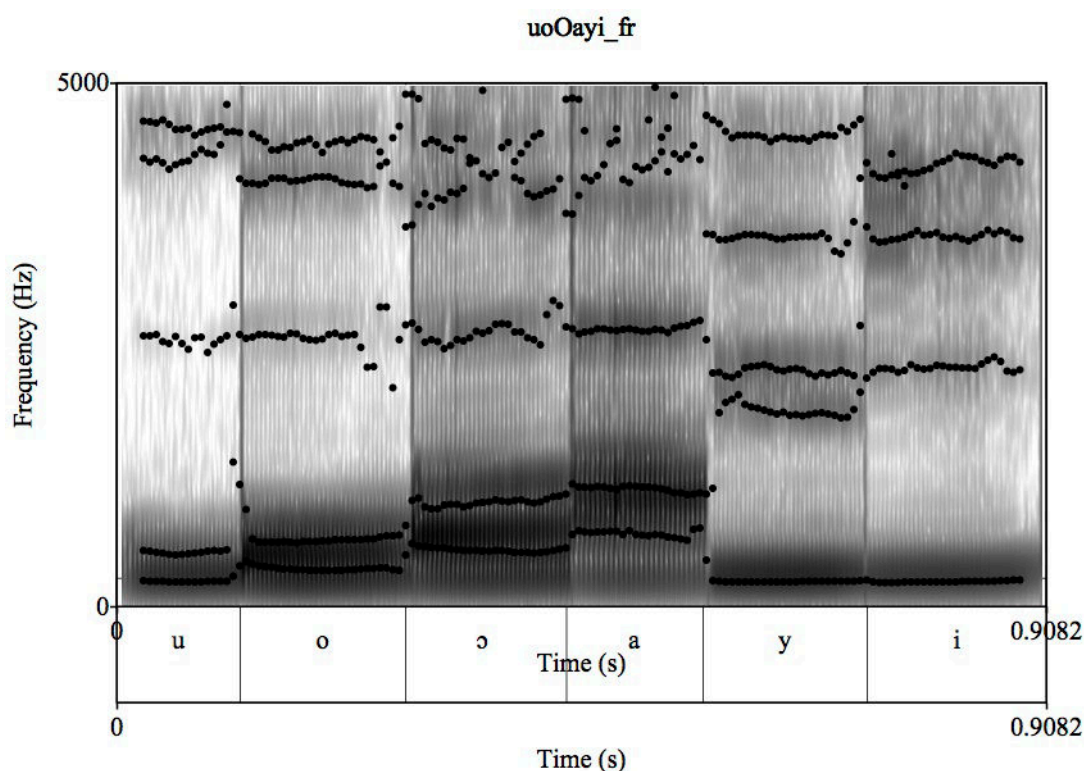


Figure 42 : Spectrogramme des voyelles du français [u, o, ɔ, a, y, i] prononcées en isolation par une Française native

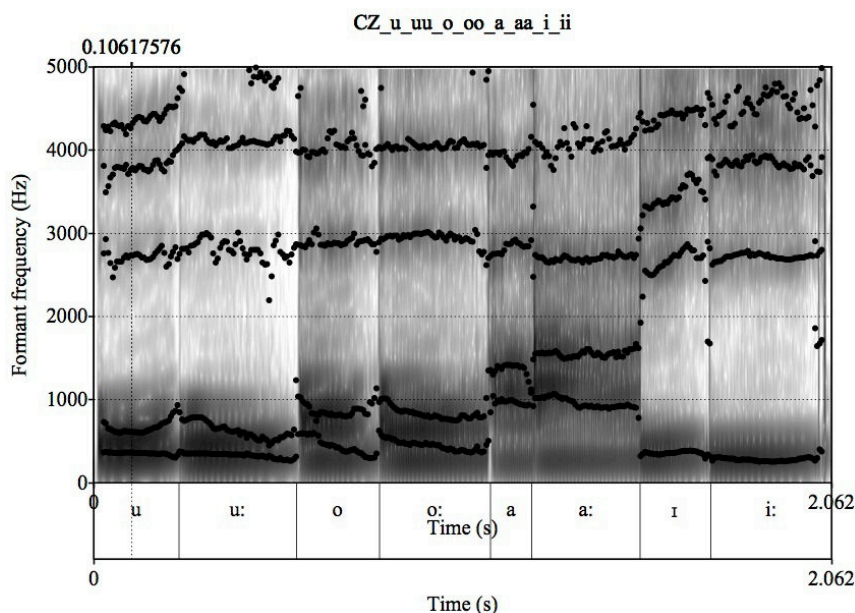


Figure 43 : Spectrogramme des voyelles du tchèque [u, u:, o, o:, a, a:, ɪ, i:] prononcées en isolation par une Tchèque native

8.3 Contrastes vocaliques

Le contraste entre les voyelles du français est assuré par le paramètre d'aperture (quatre degrés d'aperture), d'antériorité-postériorité, de labialité (opposant les voyelles étirées et arrondies) et enfin

de nasalité (opposant les voyelles orales et nasales). Le contraste des voyelles tchèques repose sur le paramètre d'aperture (trois degrés), d'antériorité-postériorité (comme en français) et de durée opposant les voyelles brèves et longues.

8.3.1 Aperture

Sur le plan de l'aperture, la valeur du premier formant permet de distinguer quatre classes des voyelles du français. Les voyelles fermées [i, y, u] se caractérisent par un premier formant bas (voir le Tableau 30) qui est respectivement de 273 Hz, 273 Hz et 288 Hz (les écarts types se trouvent entre parenthèses).

	F1	F2	F3	F4
i	273 (25, 9,2%)	2524 (271, 10,7%)	3787 (227, 6%)	4428 (195, 4,4%)
y	273 (25, 9,2%)	2037 (125, 6,1%)	2460 (190, 7,7%)	
u	288 (28, 9,2%)	770 (70, 9,1%)		
Moyenne	278			

Tableau 30 : Aperture 0 : Valeurs formantiques moyennes (en Hz) des voyelles françaises fermées (calculées à partir de 3 valeurs*10 Françaises *4 répétitions). Les écarts types (exprimés en Hz et en pourcentage) se trouvent entre parenthèses

La classe des voyelles fermées du français est également illustrée à la Figure 44 où les valeurs sont transformées en Bark et l'écart type tracé est de un.

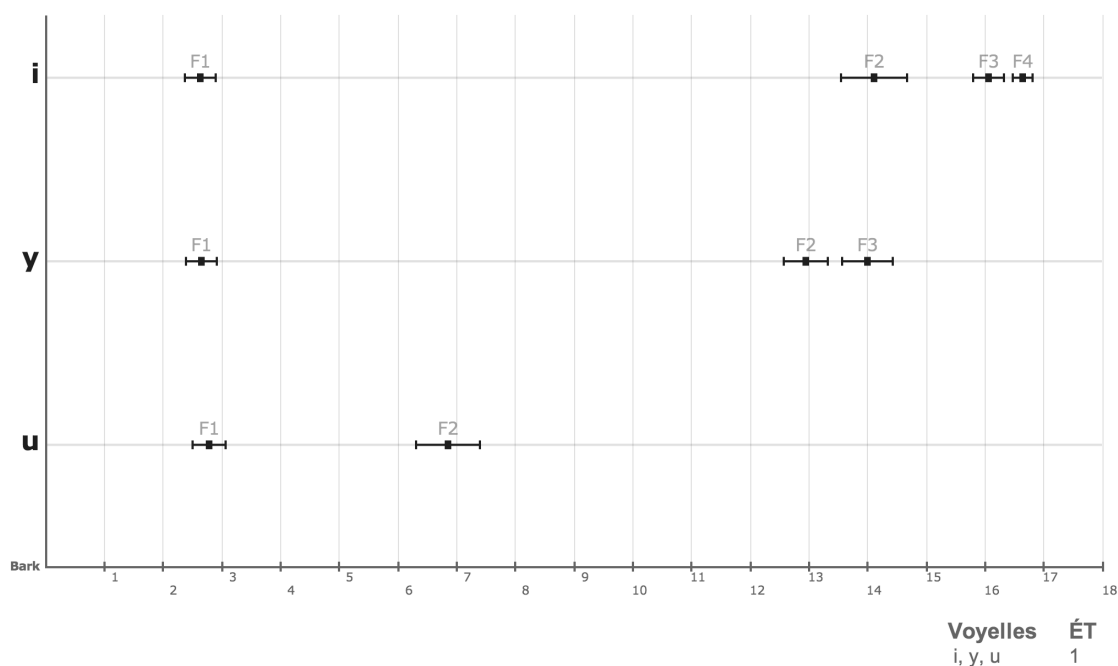


Figure 44 : Les formants F1 – F4 moyens (en Bark) de la classe des voyelles fermées du français [i, y, u]. L'écart type est de 1

Les voyelles mi-fermées [e, ø, o] se réalisent avec un premier formant plus élevé que celui des voyelles fermées, autour de 400 Hz. Les valeurs moyennes avec les écarts types entre parenthèses se trouvent dans le Tableau 31.

	F1	F2	F3
e	400 (40, 10%)	2504 (205, 8,2%)	3364 (165, 4,9%)
ø	400 (35, 8,6%)	1571 (127, 8,1%)	2645 (173, 6,5%)
o	417 (36, 8,3%)	791 (80, 10,1%)	
Moyenne	406		

Tableau 31 : Aperture 1 : Valeurs formantiques moyennes (en Hz) des voyelles françaises mi-fermées (calculées à partir de 3 valeurs*10 Françaises *4 répétitions). Les écarts types (exprimés en Hz et en pourcentage) se trouvent entre parenthèses

Les valeurs formantiques (en Bark) et les écarts types de un des voyelles mi-fermées [e, ø, o] sont également affichées à la Figure 45.

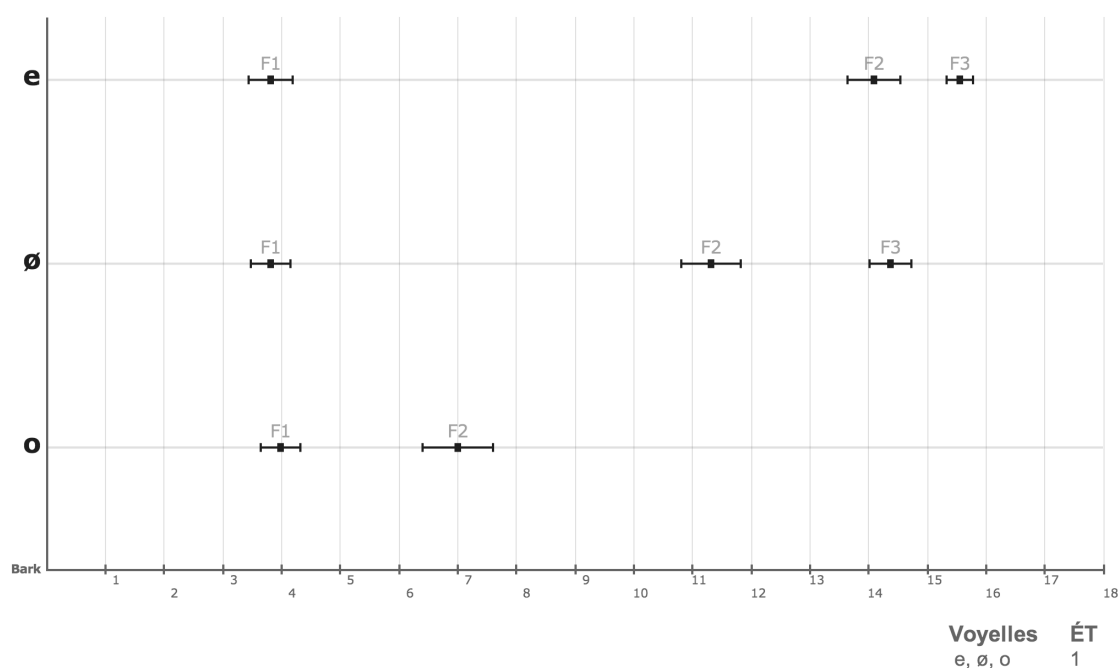


Figure 45 : Les formants F1 – F3 moyens (en Bark) de la classe des voyelles mi-fermées du français [e, ø, o]. L'écart type est de 1

La troisième classe des voyelles du français selon l'aperture correspond aux voyelles mi-ouvertes [ɛ, œ, ɔ]. Elles se réalisent avec un premier formant moyen de 624 Hz, 585 Hz et 577 Hz respectivement, comme indiqué dans le Tableau 32. Les écarts types y sont indiqués entre parenthèses.

8. Propriétés acoustiques des voyelles isolées

	F1	F2	F3
ɛ	624 (50, 8,1%)	2244 (136, 6,1%)	3066 (165, 5,4%)
œ	585 (72, 12,3%)	1579 (154, 9,8%)	2751 (229, 8,3%)
ɔ	577 (80, 13,9%)	1063 (89, 8,4%)	
Moyenne	595		

Tableau 32 : Aperture 2 : Valeurs formantiques moyennes (en Hz) des voyelles françaises mi-ouvertes (calculées à partir de 3 valeurs*10 Françaises *4 répétitions). Les écarts types (exprimés en Hz et en pourcentage) se trouvent entre parenthèses

Les formants moyens (en Bark) avec les écarts types de un des voyelles mi-ouvertes sont également illustrés à la Figure 46.

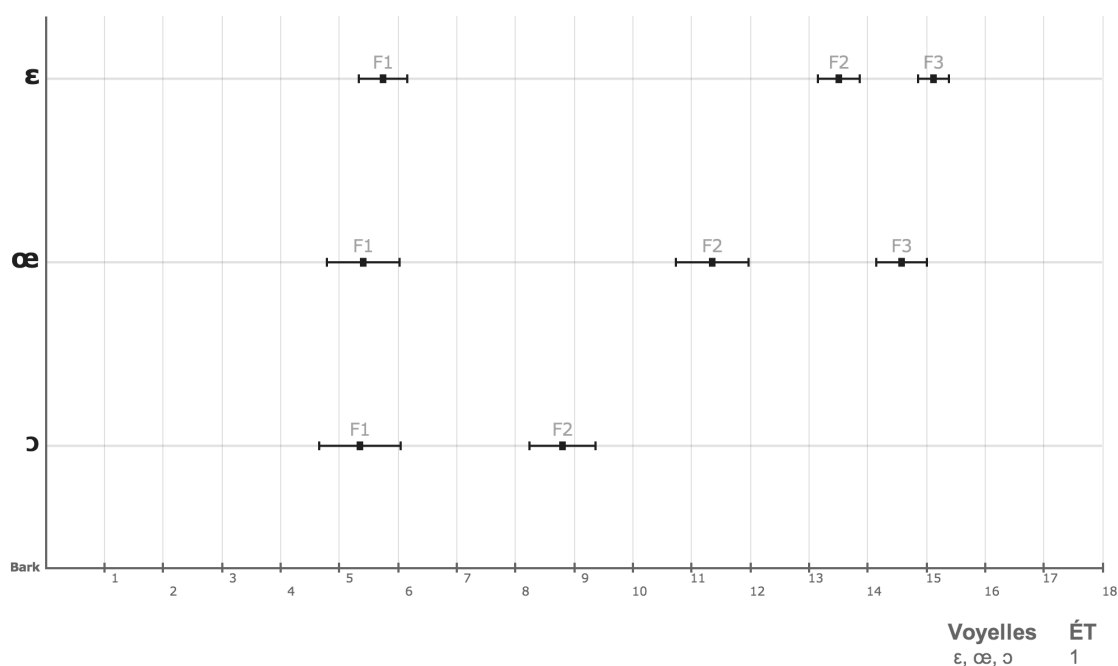


Figure 46 : Les formants F1 – F3 moyens (en Bark) de la classe des voyelles mi-ouvertes du français [ɛ, œ, ɔ]. L'écart type est de 1

Enfin, la voyelle ouverte [a] se caractérise par un premier formant maximal, qui est de 805 Hz et un écart type de 130 Hz. Son patron formantique (en Bark) est visible à la Figure 47 qui met en contraste les trois voyelles postérieures du français [u, o, ɔ] et le [a] (avec une réalisation pharyngale en isolation) ayant des degrés d'aperture différents.

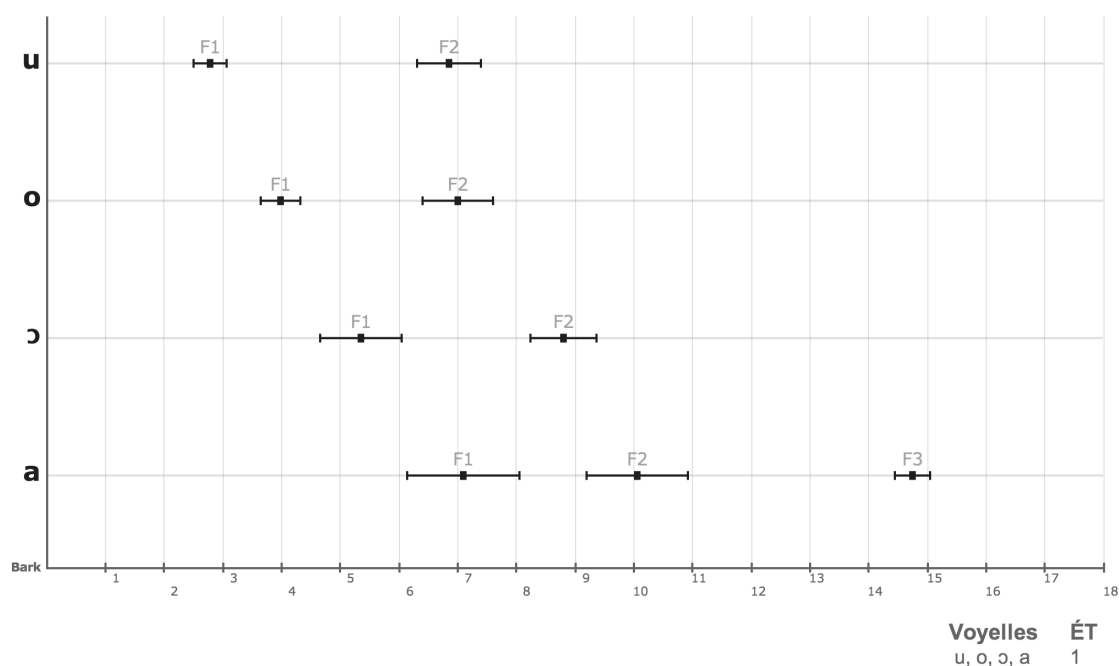


Figure 47 : Les formants F1 – F3 moyens (en Bark) des voyelles postérieures du français [u, o, ɔ,] et du [a] avec une réalisation pharyngale en isolation. L'écart type est de 1

En tchèque, les valeurs du premier formant définissent des classes moins précises. Comme indiqué dans le Tableau 33, les voyelles fermées [i:, ɪ, u:, u] se réalisent avec un premier formant moyen de 260 Hz, 403 Hz, 292 Hz, 341 Hz. Ainsi, le F1 du [ɪ] bref (403 Hz) est intermédiaire entre celui du [i:] long (260 Hz) et ceux des [ɛ] bref et [ɛ:] long qui sont respectivement de 708 Hz et 701 Hz.

	F1	F2	F3	F4
i:	260 (37, 14,2%)	2734 (153, 5,6%)	3599 (185, 5,1%)	4264 (217, 5,1%)
ɪ	403 (52, 12,9%)	2545 (133, 5,2%)	3087 (138, 4,5%)	4184 (227, 5,4%)
u:	292 (39, 13,4%)	660 (79, 12 %)		
u	341 (45, 13,2%)	718 (80, 11,1%)		
Moyenne	324			

Tableau 33 : Aperture 0 : Valeurs formantiques moyennes des voyelles tchèques fermées [i:, ɪ, u:, u] (calculées à partir de 3 valeurs*20 Tchèques *4 répétitions). Les écarts types (exprimés en Hz et en pourcentage) se trouvent entre parenthèses

La variation du premier formant (en Bark) au sein de la classe des quatre voyelles fermées du tchèque [i:, ɪ, u:, u] est également visible à la Figure 48. L'écart type affiché est de un.

8. Propriétés acoustiques des voyelles isolées

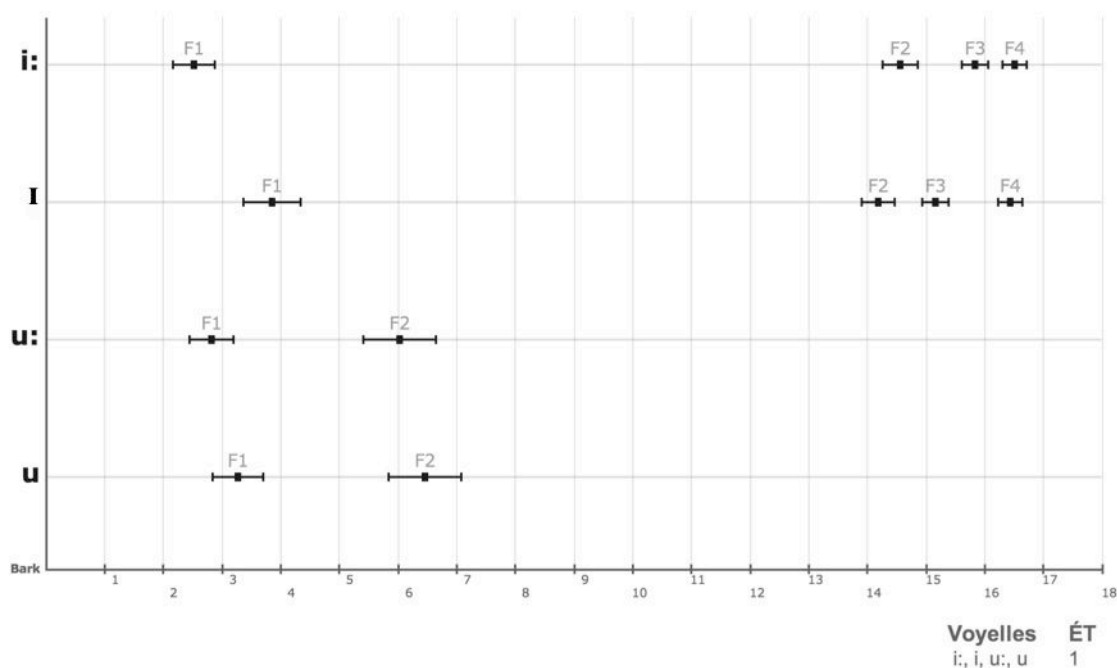


Figure 48 : Les formants F1 – F4 moyens (en Bark) des voyelles fermées du tchèque [i:, ɪ, u:, u]. L'écart type est de 1

Les voyelles d'aperture moyenne du tchèque [ɛ:, ɛ, o:, o] varient également par la valeur de leur premier formant. Alors que les voyelles antérieures [ɛ:, ɛ] se réalisent avec un F1 moyen de 701 Hz et 708 Hz respectivement, en se rapprochant davantage du [ɛ] mi-ouvert français, les voyelles postérieures [o:, o] se réalisent avec un premier formant plus bas, à savoir de 512 Hz et 540 Hz respectivement, en se situant entre le [o] mi-fermé et le [ɔ] mi-ouvert français. Le Tableau 34 donne ces valeurs moyennes de F1, F2 et F3 ainsi que les écarts types.

	F1	F2	F3
ɛ:	701 (60, 8,6%)	2094 (140, 6,7%)	2936 (155, 5,3%)
ɛ	708 (45, 6,4%)	2040 (145, 7,1%)	2913 (143, 4,9%)
o:	512 (64, 12,5%)	926 (80, 8,6%)	
o	540 (65, 12 %)	945 (94, 9,9%)	
Moyenne	615		

Tableau 34 : Aperture 1 : Valeurs formantiques moyennes (en Hz) des voyelles moyennes du tchèque [ɛ:, ɛ, o:, o] (calculées à partir de 3 valeurs*20 Tchèques *4 répétitions). Les écarts types (exprimés en Hz et en pourcentage) se trouvent entre parenthèses

Les variations du premier formant au sein de la classe des voyelles moyennes du tchèque [ɛ:, ɛ, o:, o] sont illustrées à la Figure 49. Les écarts types affichés sont de un.

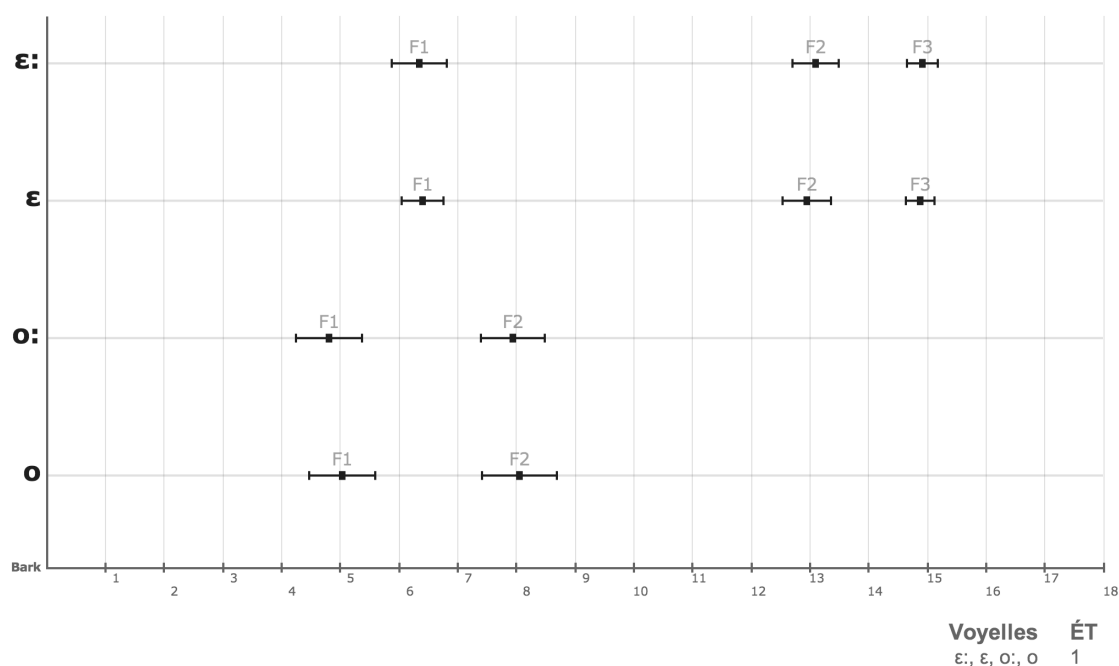


Figure 49 : Les formants F1 – F3 moyens (en Bark) des voyelles moyennes du tchèque [ɛ:, ɛ, o:, o]. L'écart type est de 1

Enfin, la troisième classe correspond aux voyelles ouvertes [a:, a] dont le premier formant moyen est respectivement de 930 Hz et 869 Hz et l'écart type de 82 Hz et 76 Hz. La Figure 50 met en contraste leurs formants moyens (en Bark) avec ceux des voyelles postérieures moyennes [o:, o] et fermées [u:, u].

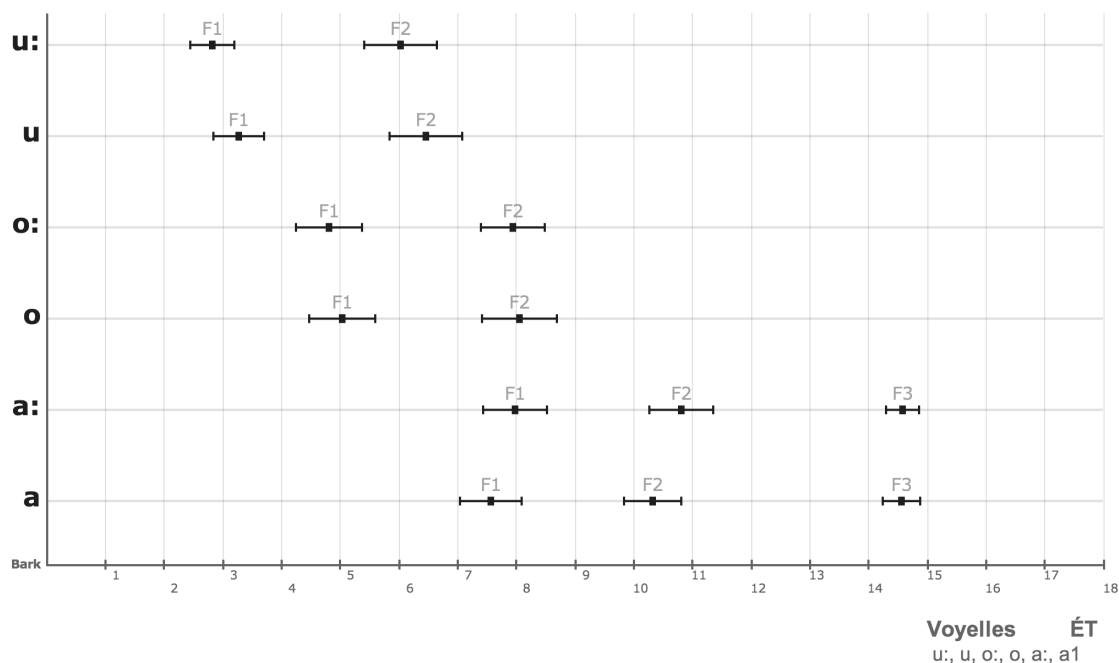


Figure 50 : Les formants F1 – F3 moyens (en Bark) des voyelles postérieures du tchèque [u:, u, o:, o] et des voyelles centrales [a:, a]. L'écart type est de 1

Ainsi, alors que le système français distingue sur la base de l'aperture quatre classes vocaliques bien précises, corrélée à la valeur du premier formant, les voyelles tchèques ont des valeurs de F1 plus variables : les voyelles fermées longues [i:, u:] ont un F1 inférieur à celui de leurs homologues brèves

[i, u] et les voyelles moyennes postérieures [o, ɔ:] ont un F1 plus bas que les voyelles moyennes antérieures [ɛ, ɛ:].

8.3.2 Antériorité/ postériorité

Comme attendu, les voyelles antérieures et postérieures du français se distinguent acoustiquement essentiellement par la valeur du deuxième formant, plus élevée pour les voyelles antérieures que postérieures. Il existe ainsi une nette séparation entre les voyelles antérieures étirées [i, e, ɛ] ayant un F2 moyen supérieur à 2000 Hz avec les voyelles antérieures arrondies [y, ø, œ] ayant un F2 supérieur à 1500 Hz et les voyelles postérieures [u, o, ɔ], ayant un F2 inférieur à 1100 Hz (pour les valeurs moyennes et les écart types exacts, se référer au Tableau 28 à la page 131). Le [a] quant à lui se réalise avec un F2 moyen d'environ 1300 Hz. Cette répartition de la valeur du deuxième formant suivant l'antériorité/ postériorité des voyelles du français est illustrée à la Figure 51 où les valeurs sont transformées en Bark et l'écart type tracé est de un. Les voyelles antérieures se trouvent dans la partie supérieure, les voyelles postérieures dans la partie inférieure.

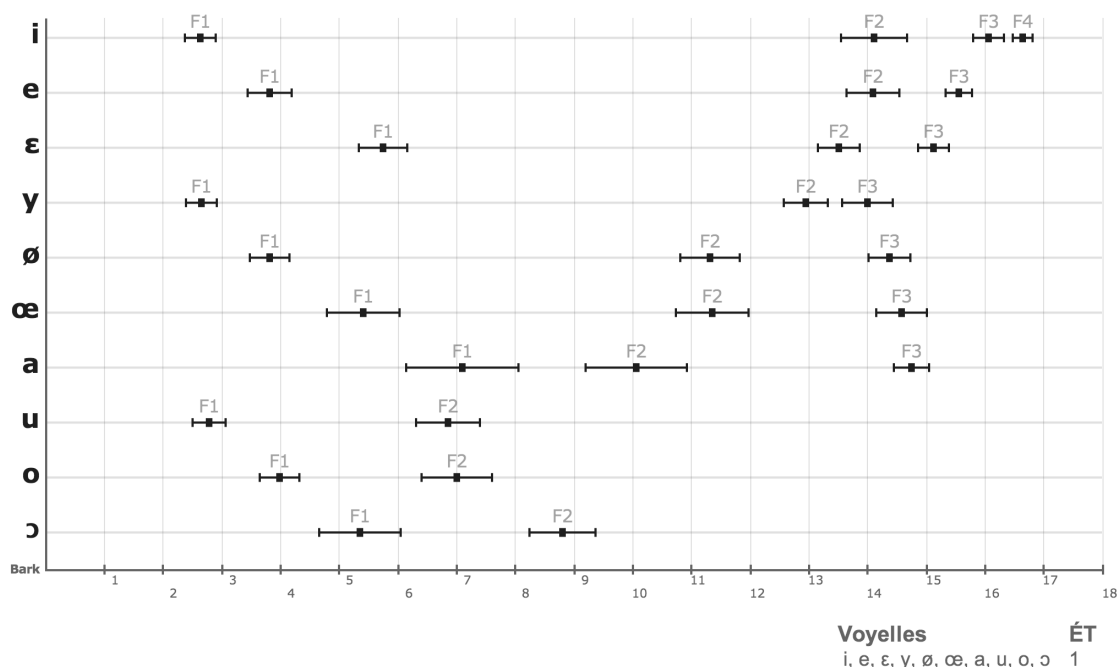


Figure 51 : Formants moyens (en Bark) des voyelles françaises antérieures (en haut) et postérieures (en bas) qui se distinguent essentiellement par la valeur du F2. L'écart type est de 1

En tchèque, l'antériorité ou la postériorité de la voyelle est essentiellement corrélée à la valeur du deuxième formant, comme attendu. Alors que les voyelles antérieures [i:, ɪ, ɛ:, ɛ] se caractérisent selon nos données par un F2 supérieur à 2000 Hz, les voyelles postérieures [u:, u, o:, o] se réalisent avec un F2 inférieur à 1000 Hz (pour les valeurs moyennes exactes et les écart types, voir le Tableau 28 à la page 131). Le [a:] long présente un F2 moyen d'environ 1450 Hz et le [a] bref un F2 d'environ 1350 Hz. La répartition de la valeur du deuxième formant en fonction de l'antériorité/ postériorité est indiquée à la Figure 52. Les valeurs y sont transformées en Bark et l'écart type affiché est de un.

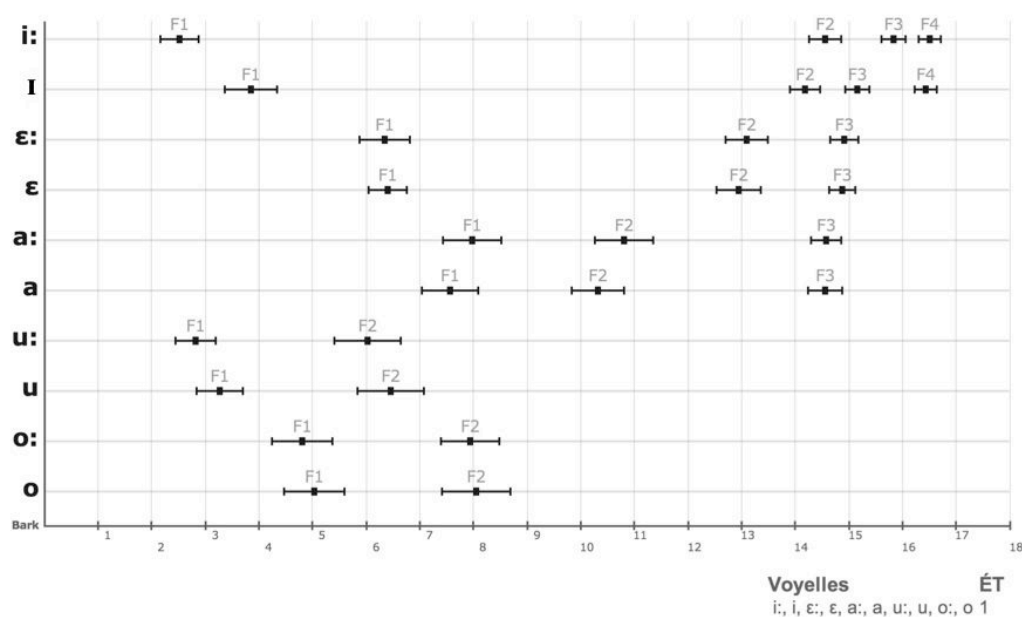


Figure 52 : Formants moyens (en Bark) des voyelles tchèques antérieures (en haut) et postérieures (en bas) qui se distinguent essentiellement par la valeur du F2. L'écart type est de 1.

En conclusion, dans les deux langues, l'antériorité/postériorité est acoustiquement corrélée essentiellement à la valeur du deuxième formant, ce dernier étant élevé pour les voyelles antérieures et bas pour les voyelles postérieures. La voyelle [a] a une réalisation pharyngale dans les deux langues, avec le F2 inférieur à 1500 Hz. La voyelle tchèque [ä:] longue est acoustiquement centrale.

8.3.3 Labialité

La labialité est un troisième paramètre utilisé pour créer des contrastes vocaliques en français : elle y oppose trois voyelles antérieures arrondies [y, ø, œ] et non-arrondies [i, e, ε]. Elle est corrélée sur le plan acoustique à la valeur des deuxième et troisième formants qui sont abaissés dans la production de voyelles arrondies par rapport à leurs homologues non-arrondies, comme illustré à la Figure 53 (avec les valeurs en Bark et écart type affiché de un).

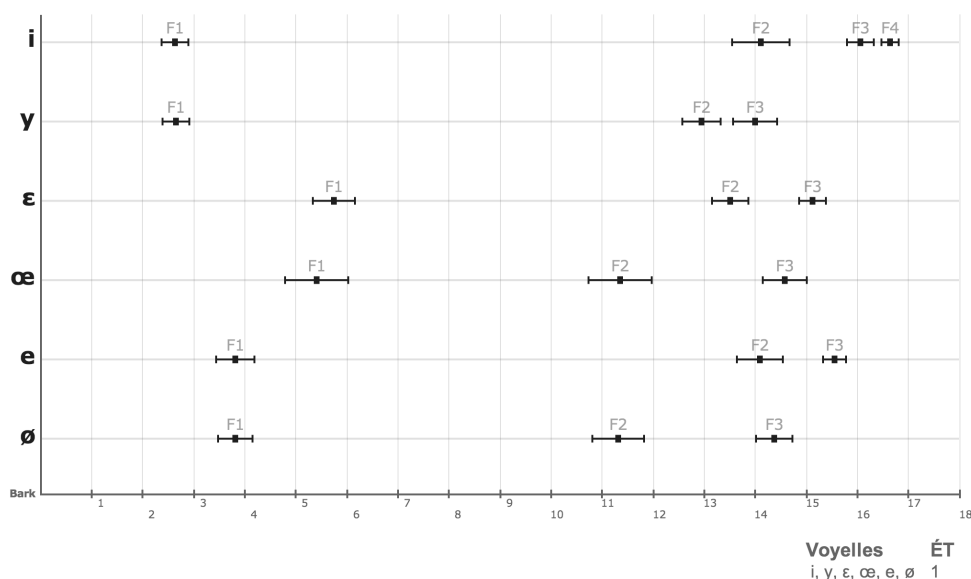


Figure 53 : Formants moyens (en Bark) des voyelles françaises antérieures labiales et étirées, ayant F2 et F3 plus élevés. L'écart type est de 1.

La Figure 53 nous fait remarquer que le couple i/y se distingue essentiellement par la valeur du troisième formant, alors que les couples e/ø et ε/œ se distinguent principalement par la valeur du deuxième formant. Il est à noter que le couple ε/œ varie légèrement par la valeur du premier formant (avec une valeur plus élevée pour [ε]) contrairement aux voyelles des deux autres couples ayant un F1 quasi identique. De cette manière, la valeur des trois premiers formants est nécessaire pour la description acoustique des voyelles antérieures, comme signalé déjà en 6.5.

8.3.4 Durée

La durée est le troisième paramètre en tchèque utilisé pour contraster les voyelles. La comparaison de la durée moyenne (en ms) des voyelles brèves et de leurs homologues longues prononcées en isolation est illustrée à la Figure 54.

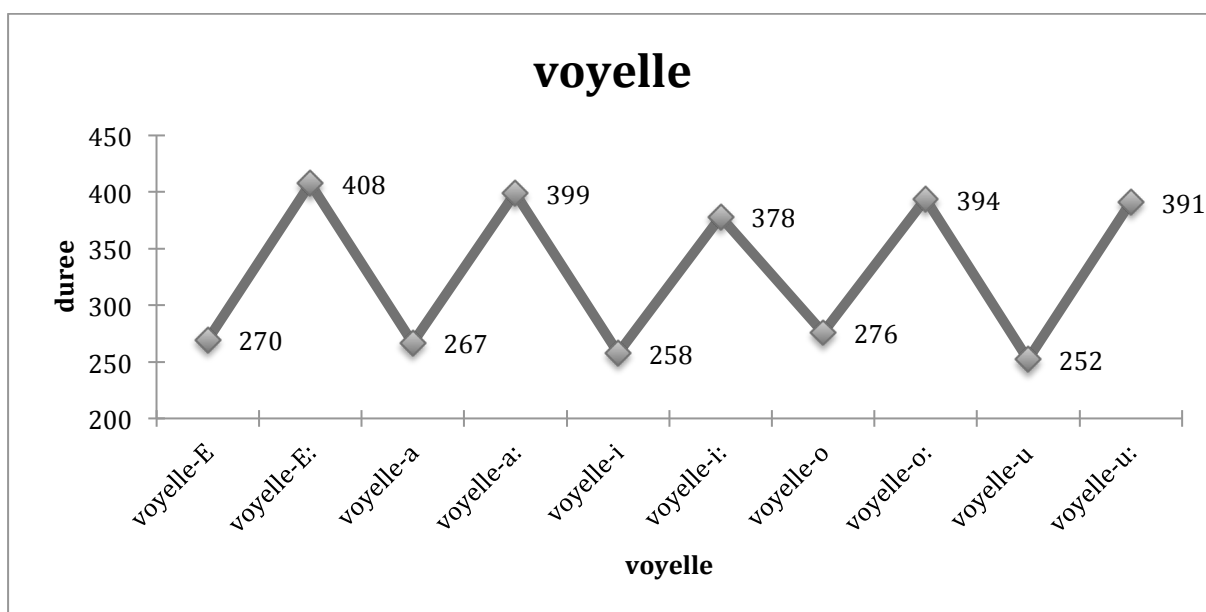


Figure 54 : Durée moyenne (en ms) des dix voyelles orales tchèques isolées [ε, ε:, a, a:, i, i:, o, o:, u, u:] chacune produite 4 fois par 20 Tchèques natives

La Figure 54 illustre que les voyelles isolées brèves sont systématiquement plus courtes que leurs homologues longues. Le rapport de durée dépend néanmoins des couples : il est de 1:1,5 dans les paires ε/ε:, a/a: , i/i: , puis de 1:1,4 dans la paire o/o: et enfin de 1:1,6 dans la paire u/u:.

Concernant les rapports de durée vocalique, la littérature traditionnelle concernant la parole spontanée (Chlumsky, 1928) indique que les voyelles longues seraient deux fois plus longues que les voyelles brèves. Des études plus récentes concluent à une différence de rapport de durée moins prononcée dans les couples i/i: et u/u: par rapport aux autres couples. Par exemple Podlipsky *et al.* (2009) reportent, à partir de lecture d'un journal par six locuteurs tchèques natifs, un rapport de durée entre i/i: nettement moins élevé (de 1,3) que celui entre les voyelles des autres paires, à savoir ε/ε:, a/a:, o/o: et u/u: qui présentent un rapport respectivement de 1 :1,7, 1 :1,8, 1 :1,7 et 1 :1,6.

Ainsi les différences entre nos résultats et ceux des études antérieures sont alors les suivantes :

- les rapports de durée pour les voyelles isolées sont moins importants que ceux pour les voyelles extraites de la parole naturelle de Chlumsky et Podlipsky, Skarnitzl et al. (2009),
- les rapports de durée sont relativement constants indépendamment des couples. Ce résultat est en accord avec les travaux classiques de Chlumsky et en désaccord avec les études récentes de Podlipsky *et al.* (2009).

Pour vérifier si les voyelles brèves et leurs homologues longues varient uniquement de par leur longueur comme l'indique la littérature traditionnelle ou également de par leurs caractéristiques spectrales, comme le montrent les études plus récentes (Podlipsky *et al.*, 2009; Skarnitzl and Volin, 2012), nous avons comparé également les valeurs de leurs formants. Comme le montre la Figure 55, les voyelles des paires $\epsilon:/\epsilon$, $o:/o$, $a:/a$ rencontrent des valeurs formantiques semblables car elles se situent à moins d'un écart type les unes des autres (pour les valeurs moyennes ainsi que les écarts types précis, voir le Tableau 28 à la page 131).

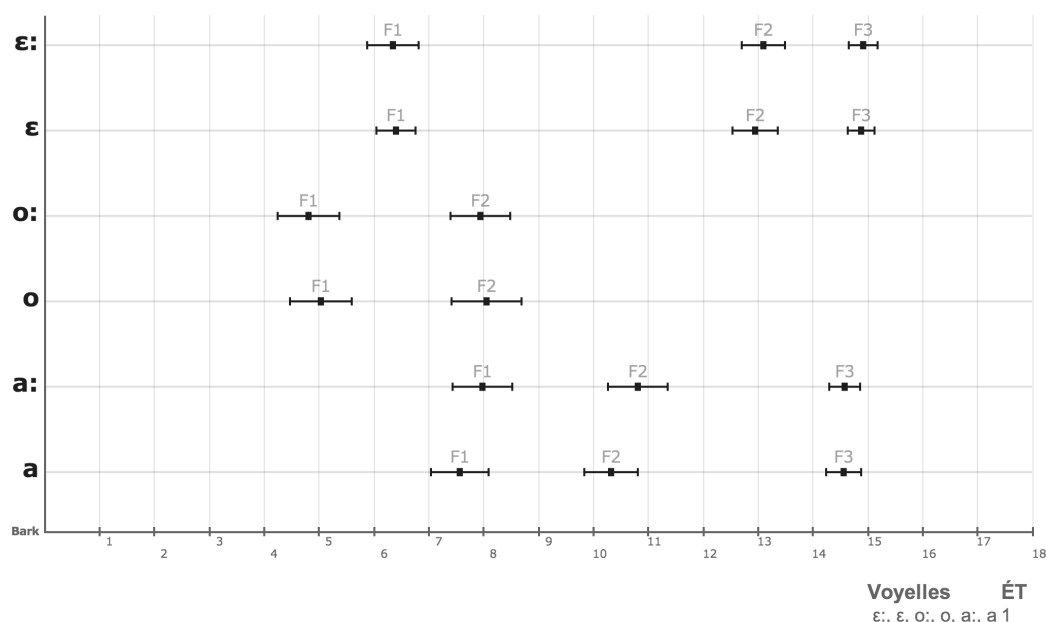


Figure 55 : Formants moyens (en Bark) des voyelles tchèques brèves [ϵ , o , a] et de leurs homologues longues [$\epsilon:$, $o:$, $a:$]. L'écart type est de 1

Ainsi, le contraste entre les voyelles isolées des trois couples $\epsilon:/\epsilon$, $o:/o$, $a:/a$ repose entièrement sur le paramètre de durée.

En revanche, les voyelles des paires $i:/i$ (et $u:/u$) ne varient pas uniquement de par leur durée mais également de par la valeur de leurs formants comme visible à la Figure 56.

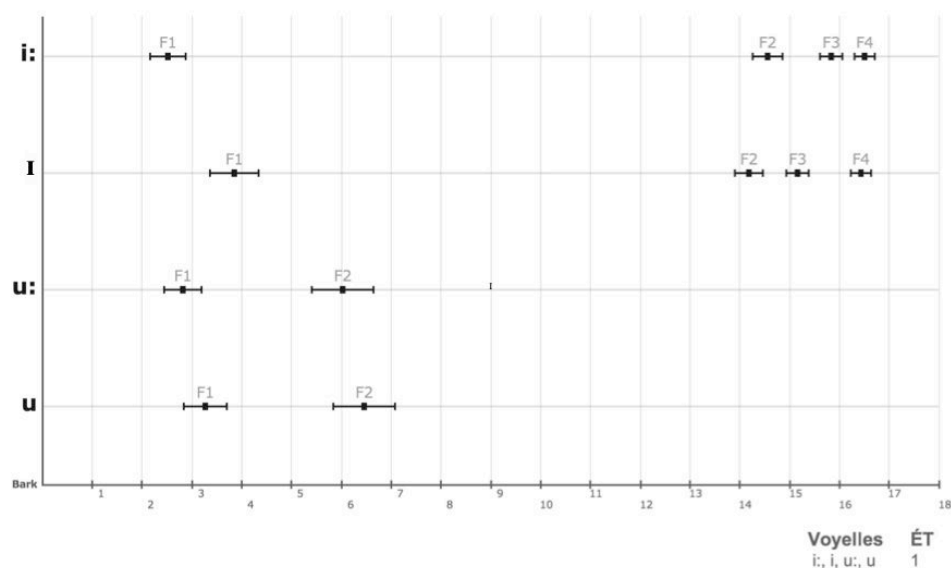


Figure 56 : Formants moyens (en Bark) des voyelles tchèques brèves [i , u] et de leurs homologues longues [$i:$, $u:$]. L'écart type est de 1

8. Propriétés acoustiques des voyelles isolées

Au niveau de la paire i:/ɪ, la différence du patron formantique est évidente car les valeurs varient au niveau des formants F1, F2 et F3. Ainsi, le [i:] long se caractérise par un F1 plus bas et des F2 et F3 plus élevés que le [ɪ] bref. À l'inverse, la différence acoustique entre le [u] bref et son homologue long est moins marquée. Elle concerne essentiellement la valeur du F1 dont la distance dépasse légèrement un écart type. De cette façon, le [u:] long se réalise avec un F1 plus bas que le [u] bref.

Pour pouvoir comparer nos données avec les données de Skarnitzl and Volin (2012) sur les différences de timbre (l'étude est détaillée en 2.2.1.3.2 : valeurs formantiques des voyelles issues de la lecture par 23 locuteurs masculins) entre les voyelles brèves et leurs homologues longues, nous avons effectué une transformation des valeurs en unités ERB et nous avons calculé la différence (en pourcentage) entre les valeurs formantiques au niveau de chaque paire. Le résultat est indiqué dans le Tableau 35 à la page 148 qui renseigne les valeurs moyennes des formants en Hz, en unités ERB et indique la différence entre les formants au niveau de chaque couple vocalique. Les écarts types se trouvent entre parenthèses.

Nous remarquons que comme dans l'étude de Skarnitzl et Volin, tous les formants vocaliques comparés des paires ε:/ε, o:/o, a:/a ont une différence de valeurs inférieure à 5 %, ce qui soutient l'idée que ces voyelles possèdent le même timbre. En revanche, la différence du premier formant des voyelles de la paire i:/ɪ est de 30,6 % chez Skarnitzl et Volin et de 36,9 % dans nos données. La différence de F2 du couple i:/ɪ indiquée par les auteurs est de 5,8 % alors que nos résultats montrent une différence moindre, à savoir de 2,8 %. Il existe également une différence du premier formant des voyelles du couple u:/u qui est néanmoins moins importante que celle entre i:/ɪ : nos données montrent une différence de 11,3 % et selon Skarnitzl et Volin elle est de 12,6 %. Ces derniers observent également une différence au niveau de F2 entre le [u] bref et le [u:] long produits en contexte, dans la parole lue, qui est de 11,4 % alors que nous ne notons une différence que de 5 % pour ces voyelles produites en isolation.

Les auteurs relient les différences spectrales aux différents rapports de durée entre les voyelles brèves et leurs homologues longues. Le rapport de durée est, selon les auteurs, de 1 :1,7 à 1,8 au sein des paires ε:/ε, o:/o, a:/a, de 1,6 dans le cas de u:/u et de 1 :1,3 dans le cas de i:/ɪ. Les paires de voyelles avec des différences spectrales plus importantes seraient ainsi accompagnées de différence de durée moins marquée. Notons que nos données ne soutiennent pas ce postulat, du moins en ce qui concerne les voyelles produites en isolation. Comme précisé ci-dessus, la paire i:/ɪ rencontre le même rapport de durée que a:/a ou ε:/ε (qui est de 1 :1,5) et la paire u:/u rencontre un rapport de durée le plus important des cinq couples à savoir de 1 :1,6.

Nos données acoustiques sur les voyelles isolées du tchèque montrent que les différences spectrales entre i:/ɪ et éventuellement u:/u renforcent (mais ne suppléent pas) la différence de durée.

Pour vérifier si la différence des formants F1 (de 11,3 % selon nos données) et F2 (de 5 %) des voyelles u:/u de même que la différence formantique de F1 et F3 des voyelles i:/ɪ est perçue et à l'inverse que les couples ε:/ε, o:/o, a:/a correspondent à seulement trois timbres auprès des auditeurs tchèques natifs, nous avons conduit une expérience perceptive. Nous avons alors raccourci toutes les dix voyelles monophthongues du tchèque (répétées au total onze fois par cinq locutrices tchèques natives de la région de Bohême) à 150 ms en gardant la partie centrale de la voyelle. La description de l'expérience et des résultats se trouve dans les annexes (page 44 de l'annexe du chapitre 8).

Le résultat du test d'identification de voyelles monophthongues abrégées à 150 ms montre que les voyelles longues [a:], [ε:], [o:] et [u:] sont difficilement identifiées comme telles sur la seule base du timbre : lorsqu'elles sont abrégées, leur catégorisation se fait majoritairement en termes de leur homologue brève, comme attendu. En revanche, la voyelle longue [i:] abrégée à 150 ms n'est jamais confondue par les auditeurs natifs avec son homologue brève. Ainsi, les voyelles du couple i:/ɪ ne se distinguent pas seulement par la durée, comme les voyelles des autres couples, mais également par le timbre vocalique.

	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)	F1 (ERB)	F2 (ERB)	F3 (ERB)	F4 (ERB)	Diff F1 %	Diff F2 %	Diff F3 %	Diff F4 %
i	403 (52)	2545 (133)	3087 (138)	4184 (227)	8,9 (0,8)	22,9 (0,4)	24,5 (0,4)	27 (0,4)				
i:	260 (37)	2734 (153)	3599 (185)	4264 (217)	6,5 (0,7)	23,5 (0,5)	25,8 (0,4)	27,1 (0,4)	36,9	2,6	5,1	0,4
ε	708 (45)	2040 (145)	2913 (143)		12,7 (0,5)	21,1 (0,6)	24 (0,4)					
ε:	701 (60)	2094 (140)	2936 (155)		12,6 (0,6)	21,3 (0,6)	24,1 (0,4)		0,8	0,9	0,4	
a	869 (76)	1344 (101)	2726 (164)		14,2 (0,7)	17,6 (0,6)	23,5 (0,5)					
a:	930 (82)	1452 (120)	2739 (152)		14,7 (0,7)	18,2 (0,7)	23,5 (0,5)		3,4	3,3	0	
o	540 (65)	945 (94)			10,8 (0,8)	14,8 (0,8)						
o:	512 (64)	926 (80)			10,4 (0,8)	14,7 (0,7)			3,8	0,7		
u	341 (45)	718 (80)			7,9 (0,7)	12,7 (0,8)						
u:	292 (39)	660 (79)	2751 (229)		7,1 (0,7)	12,1 (0,8)			11,3	5		

Tableau 35 : Les valeurs formantiques des voyelles tchèques, en Hz (1^e colonne), en ERB (2^e colonne) et la différence en pourcentage (3^e colonne) entre les valeurs formantiques des voyelles brèves et longues au niveau de chaque paire. Les écarts types sont entre parenthèses

8.4 Similarité acoustique des voyelles isolées du français et du tchèque

Puisque le SLM (Flege, 1995) prédit une performance native pour les voyelles *identiques*, nous voulons tout d'abord tester s'il existe des voyelles véritablement *identiques* sur le plan acoustique, c'est à dire avec une différence non significative des valeurs formantiques ($p > 0,05$) entre le français et le tchèque. Nous considérons néanmoins les résultats des tests statistiques avec précaution, car comme le précise Flege (1987b, p. 291) : « *It is, of course, difficult to make valid, cross-language, comparisons of parameter values for a phonetic dimension.* ».

Nous émettons l'hypothèse qu'il existe des voyelles isolées acoustiquement *identiques* entre le français et le tchèque. Autrement dit, nous stipulons que certaines voyelles isolées du français et du tchèque ne diffèrent pas significativement par la valeur des formants comparés (F1 et F2 des voyelles postérieures, F1, F2 et F3 des voyelles antérieures, F1, F2, F3 et F4 du [i], [ɪ] ou [i:]).

L'hypothèse nulle consiste à dire qu'il n'existe pas de voyelles acoustiquement identiques entre le français et le tchèque.

Pour vérifier notre hypothèse à savoir qu'il existe des voyelles isolées acoustiquement *identiques* entre le français et le tchèque, nous avons comparé les formants un par un des voyelles acoustiquement proches dans un test-t bilatéral. Les paires des voyelles acoustiquement proches que nous avons comparées sont les suivantes : [i]fr - [i:]tch, [e]fr - [ɪ] tch, [ɛ]fr - [ɛ]tch, [ɛ]fr - [ɛ:]tch, [a]fr - [a]tch, [o]fr - [o]tch, [o]fr - [u]tch, [ɔ]fr - [o]tch, [ɔ]fr - [a]tch, [u]fr - [u:]tch, [œ]fr - [ɛ]tch.

Les deux premiers formants des voyelles postérieures, les trois premiers formants des voyelles antérieures et les quatre premiers formants de [i, ɪ, i:] ont été comparés. Le résultat du test-t fournit la valeur $p < 0,05$ au niveau de tous les formants des voyelles comparées, mis à part pour le F1 des paires [e]fr - [ɪ]tch où $p = 0,527$ et [u]fr - [u:]tch où $p = 0,226$. Le résultat intégral du test-t bilatéral se trouve en annexe du chapitre 8, à la page 44 des annexes.

Les apprenants tchèques n'ont donc jamais affaire à des sons acoustiquement *identiques* dans l'apprentissage des voyelles isolées du français car il existe des différences significatives entre les moyennes d'au moins un formant comparé. Notre hypothèse de départ est alors rejetée et l'hypothèse nulle est retenue.

Dans l'impossibilité d'utiliser le terme voyelles *identiques*, nous opposerons alors les voyelles *similaires* aux voyelles "hautement similaires" dont les formants se situent à un écart type ou moins des moyennes formantiques des voyelles tchèques.

Comme exposé dans la partie méthodologique du chapitre 6, la similarité acoustique des voyelles du français et du tchèque dépend des distances de leurs formants et nous les définissons de la manière suivante :

Les voyelles du français "hautement similaires", avec *similarité* +++, sont celles dont les formants moyens se trouvent à un écart type ou moins des formants moyens des voyelles tchèques.

Les voyelles du français *similaires* avec *similarité* ++ sont celles dont les formants se trouvent entre un et 1,5 écart type des formants des voyelles tchèques.

Les voyelles du français *similaires* avec *similarité* + sont celles dont les formants se trouvent entre 1,5 et deux écarts types des formants des voyelles tchèques.

Les voyelles du français *nouvelles* avec *similarité* 0 sont celles dont un ou plusieurs formants se trouvent à plus de deux écarts types des formants des voyelles tchèques.

La Figure 57 représente le triangle F1/F2 (l'image du haut) et le triangle F2/F3 (l'image du bas) des voyelles orales du français (en pointillé) et des voyelles monophthongues du tchèque (en trait plein). L'écart type affiché est de 1 et l'échelle est en Bark. Observons la distance acoustique entre les voyelles du français et du tchèque.

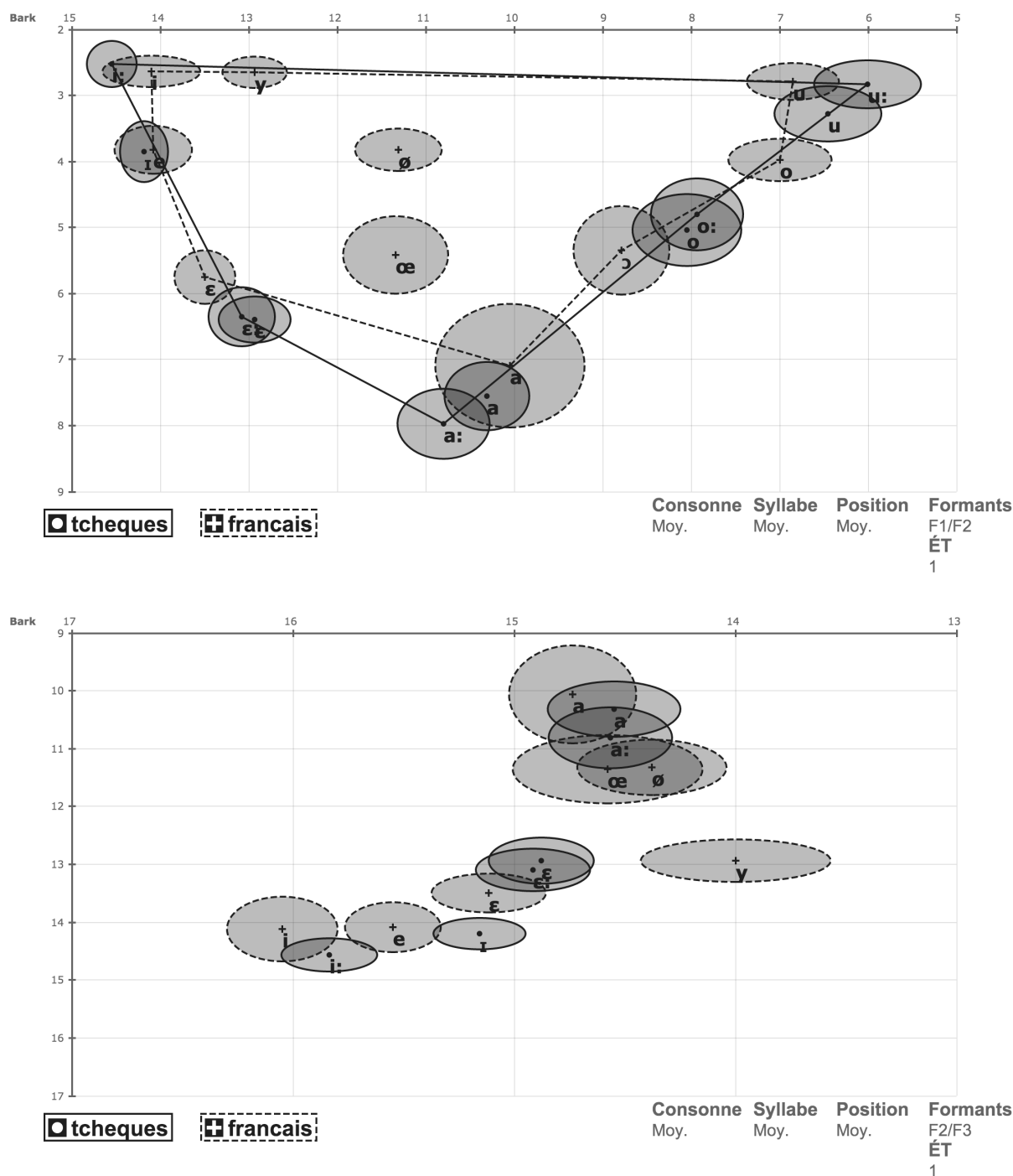


Figure 57 : Triangle vocalique (en Bark) des voyelles orales isolées du français (en pointillé) et du tchèque (en trait plein) sur le plan F1/F2 (en haut) et F2/F3 (en bas). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 10 Françaises/ 20 Tchèques*4 répétitions*3 valeurs par voyelle

La Figure 57 montre qu'il n'y a qu'une voyelle du français qui soit « hautement similaire » par tous ses formants comparés : il s'agit de [a] dont les trois premiers formants se trouvent à un écart type ou moins de ceux du [a] tchèque. Nous remarquons que même si cette voyelle est « hautement similaire », elle est plus dispersée que le [a] tchèque. Les triangles vocaliques F1/F2 et F2/F3 avec des ellipses de dispersion tracées à 1,5 et deux écarts types se trouvent en annexe du chapitre 8 (pages 48 – 50 des annexes).

Le Tableau 36 résume la similarité acoustique entre les voyelles du français (FR) et des voyelles du tchèque (TCH) du point de vue des écarts types de leurs moyennes formantiques.

8. Propriétés acoustiques des voyelles isolées

+ - Similarité acoustique contexte 0	FR	TCH	≤ 1 ET	1 - 1,5 ET	1,5 - 2 ET	> 2 ET
	<u>a</u>	<u>a</u>	F1/F2 F2/F3			
	ɔ	o		F1/F2		
	u	u/ u:		F1/F2		
	ɛ	ɛ			F1/F2 F2/F3	
	o	u			F1/F2	
	e	ɪ	F1/F2		F2/F3	
	i	i:		F1/F2	F2/F3	
	y	/				✓
	ø	/				✓
	œ	/				✓

Tableau 36 : Similarité acoustique entre les voyelles du français (FR) et du tchèque (TCH) prononcées en isolation, exprimée en termes d'écart types (ET). <= 1 ET => Similarité +++ (gris foncé), 1 – 1,5 ET => Similarité ++ (gris moyen), 1,5 – 2 ET => Similarité + (gris clair), 2 et + ET => Similarité 0

Le Tableau 36 met en évidence que :

- [a] isolé est par ses formants « hautement similaire » (avec similarité +++) au [a] tchèque
- [y, ø, œ] isolés sont des voyelles acoustiquement *nouvelles* (similarité 0)
- [ɔ, u] sont par leurs formants *similaires* (avec similarité ++) à respectivement [o, u/u:] tchèques
- [ɛ, o, e, i] sont par leurs formants *similaires* (avec similarité +) à respectivement [ɛ, u, ɪ, i:] tchèques

La Figure 58 de la page 152 est un spectrogramme illustrant la similarité des voyelles du français (fr), prononcées par une Française native, et du tchèque (tch), prononcées par une Tchèque native. Nous comparons les formants F1, F2 des voyelles postérieures, F1, F2 et F3 des voyelles antérieures et enfin F1, F2, F3 et F4 de [i, ɪ, i:].

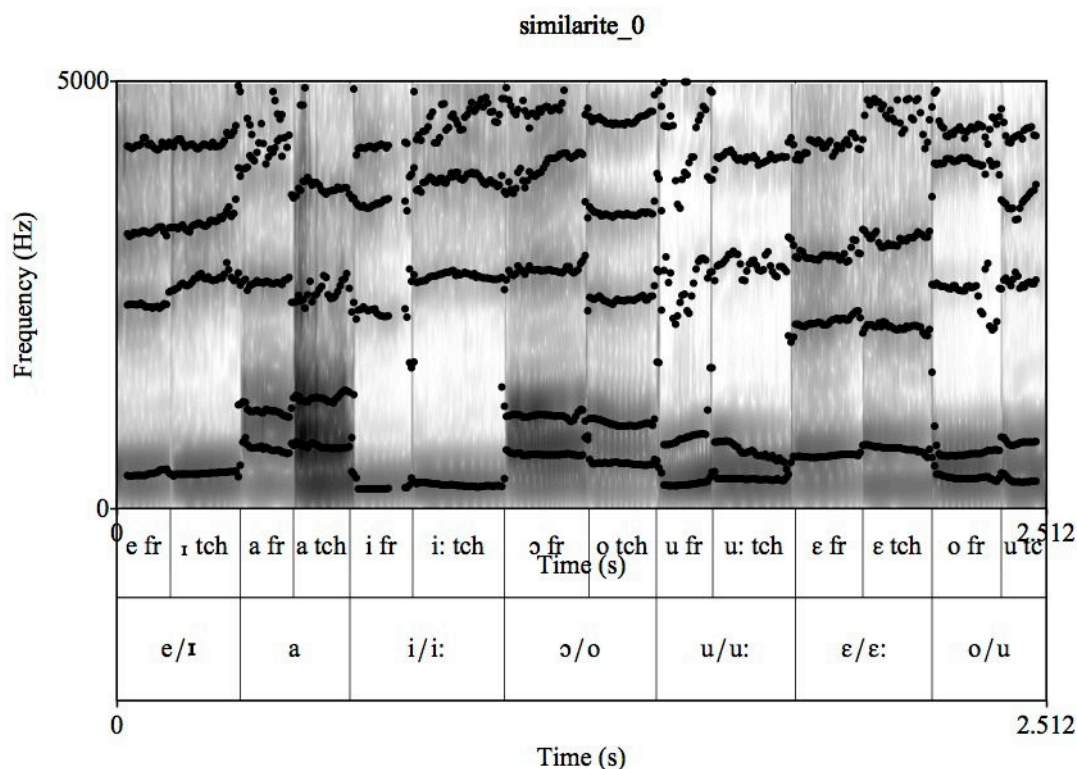


Figure 58 : Spectrogramme des paires de voyelles du français et du tchèque [e-ɪ], [a-a], [i-i:], [ɔ-o], [u-u:], [ɛ-ɛ], [o-u] prononcées par une Française et une Tchèque. Notons la similarité des patrons formantiques

Comme l'indique la Figure 58, une certaine similarité formantique peut être observée entre les paires des voyelles suivantes : [e-ɪ], [a-a], [i-i:], [ɔ-o], [u-u:], [ɛ-ɛ], [o-u] où le premier membre de chaque paire est une voyelle française et le deuxième membre une voyelle tchèque.

Pour résumer, les voyelles isolées du français acoustiquement *identiques* aux voyelles tchèques n'existent pas mais [a] est par ses formants « hautement similaire » (avec similarité +++) au [a] tchèque. Les voyelles [y, ø, œ] sont des voyelles acoustiquement *nouvelles* (similarité 0). Les autres voyelles [ɔ, u, ɛ, o, e, i] sont acoustiquement *similaires* (avec similarité ++ ou +) aux voyelles tchèques. Il reste à vérifier la similarité perceptive de ces voyelles avec les voyelles tchèques ce qui fait l'objet du chapitre 11.

9. Propriétés acoustiques des voyelles en contextes phonétiques p, t, k, R/h : Effets de la coarticulation

Les sons du français ne se contentent pas d'être des phonèmes isolés. A l'instar des individus, ils n'existent que par et pour les autres : ils vivent en société.

(Wioland, 2005, p. 26)

Résumé : Dans le **chapitre neuf** nous comparons les effets des consonnes environnantes sur les formants et la durée vocalique en français et en tchèque. Les prédictions sur la coarticulation sont dégagées à partir de la synthèse articulatoire par le modèle Maeda et le sens ainsi que l'ampleur des variations acoustiques effectives sont définis par rapport aux valeurs formantiques des voyelles isolées (la cible). Le résultat montre que le sens des mouvements formantiques de toutes les voyelles, à l'exception du [ɔ] français, s'explique soit par les loci des consonnes environnantes, soit par la centralisation acoustique. L'élévation inattendue du F1 du [ɔ] en contexte labial par rapport à la cible peut s'expliquer par l'enfreint des règles orthoépiques lors de la production d'un [ɔ] isolé.

Les deux langues diffèrent ensuite par l'ampleur des variations acoustiques des voyelles antérieures et postérieures. En français, le F2 des voyelles antérieures (qui composent deux séries : étirées et arrondies) varie dans une moindre mesure qu'en tchèque (qui comporte seulement les voyelles étirées). Ce résultat est attendu (Manuel and Krakow, 1984). A l'inverse, le F2 des voyelles postérieures varie moins en tchèque qu'en français où une grande antériorisation acoustique des voyelles postérieures en contexte consonantique antérieur a déjà été observée par exemple par De Mareüil *et al.* (2010). Les différents effets de coarticulation permettent d'expliquer pourquoi le degré de similarité acoustique des voyelles des deux langues change d'un contexte à l'autre. Nous avons mis en évidence que selon le contexte dans lequel les voyelles avaient été prononcées, la même voyelle du français (par exemple le [ɔ]) peut être soit acoustiquement *nouvelle* (en contexte labial, dental ou palato-vélaire), soit acoustiquement *similaire* à une voyelle tchèque (en isolation aux [o/o:]). Cependant aucune voyelle du français n'est acoustiquement *identique* à une voyelle du tchèque. Ces résultats ont une implication pour les modèles d'apprentissage phonétique (comme le SLM) dont les hypothèses se basent sur la notion de 'similarité'.

La réduction vocalique due à la coarticulation se manifeste essentiellement par une modification du patron formantique et par un abrègement de la voyelle, ce qui est typiquement l'effet des consonnes obstruantes sourdes en coda dans les syllabes fermées (Flemming, 2006). Une définition de la réduction vocalique classique, quoique critiquable comme exposé dans le paragraphe suivant, est celle de Miller (1981, p. 42) :

Vowel reduction refers to the tendency for the obtained formant frequencies of a vowel to fall short of the idealized target values for that vowel – those values that would be obtained if the vowel were produced in isolation – resulting in an overall shrinkage of the vowel space.

Alors que traditionnellement on postule un mécanisme de centralisation systématique des valeurs formantiques qui tendent vers le schwa (Delattre, 1969), Lindblom (1963) et Stevens and House (1963) expliquent les déplacements formantiques essentiellement par un mécanisme d'assimilation contextuelle, c'est-à-dire des mouvements formantiques en direction des loci des consonnes environnantes. Notons que l'assimilation contextuelle ne va pas toujours dans la même

direction qu'une centralisation acoustique. Par exemple, les consonnes produites avec une constriction postérieure comme l'uvulaire [ɣ] provoquent une élévation du premier formant (Stevens, 1998). Ainsi, le F1 du [ɔ] français (mi-ouvert) est plus élevé dans le contexte [ɣ] qu'à l'isolée (+ 77 Hz) selon nos données (tableau récapitulé en annexe du chapitre 9, page 60 des annexes), ce qui est une observation en faveur du mécanisme de coarticulation.

Même si la réduction phonétique est un phénomène universel, son ampleur varie d'une langue à l'autre et l'apprenant doit apprendre à coarticuler les voyelles selon les motifs de coarticulation de la LE.

Concernant la coarticulation des voyelles du français et du tchèque dans les mêmes environnements consonantiques (labial, dental et palato-vélaire), nous nous sommes posée les trois questions suivantes:

1) Quels sont les effets acoustiques des contextes symétriques labial, dental et palato-vélaire sur les voyelles orales du français et du tchèque ? Le degré de coarticulation varie-t-il en fonction de la langue? Selon l'hypothèse de Manuel (1990), le français ayant un nombre de qualités vocaliques plus élevé (dix qualités vocaliques en français actuel), devrait présenter un degré de coarticulation moindre que le tchèque, ayant un nombre de timbres réduit (traditionnellement cinq, actuellement six ou sept selon les études).

2) Dans quelle mesure le contexte consonantique influence-t-il le caractère focal des voyelles (qui sont focales en isolation)? L'atténuation du caractère focal des voyelles en contexte par rapport à la cible, si elle se produit, est-elle de la même ampleur dans les deux langues ?

3) Enfin, comment varie la similarité acoustique des voyelles des deux langues en fonction du contexte consonantique ?

Pour répondre à la première question, nous avons calculé la distance acoustique entre les formants des voyelles en contextes [p, t, k] et ceux des voyelles en isolation, qui définissent en principe la cible acoustique (Miller, 1981), ou tout au moins peuvent servir de référence. Les valeurs formantiques des voyelles en contexte ont été relevées au milieu de la voyelle (dans la partie stable où la voyelle se rapproche le plus de la cible) alors que les valeurs formantiques des voyelles isolées (la cible) correspondent à la moyenne des valeurs mesurées à un tiers, à la moitié et à deux tiers de la durée vocalique. Plus la distance acoustique entre la valeur des formants des voyelles en contexte et celle des formants des voyelles en isolation augmente, plus la variation due au contexte est grande. Les changements formantiques peuvent être positifs, correspondant à une élévation de la fréquence du formant, ou négatifs, selon la nature de la consonne et de la voyelle subséquente.

Pour répondre à la deuxième question, nous avons comparé les distances acoustiques F2-F1 (des voyelles postérieures), F3-F2 (des voyelles antérieures) et F4-F3 (de [i, i:]) des voyelles en contexte (valeurs prises au milieu de la voyelle) et celles des voyelles en isolation (moyenne des valeurs prises à un tiers, à la moitié et à deux tiers de la durée vocalique).

Enfin, pour répondre à la troisième question, nous avons comparé les voyelles orales du français et du tchèque en contextes [p, t, k] en termes de moyennes et de leurs écarts types et nous avons mis en contraste les résultats avec ceux concernant les voyelles isolées.

9.1 Prédictions à partir d'un modèle articulatoire (modèle Maeda)

Dans un premier temps, nous avons établi des prédictions sur les changements formantiques des voyelles extrêmes [i, ɑ, u] et de la voyelle labiale [y] dus au contexte consonantique en utilisant VTDemo (Huckvale, 2009) qui est un programme de synthèse articulatoire décrit en 4.1.3.

L'expérience et les résultats montrant les spectres, les spectrogrammes et les conduits vocaux des sons résultants se trouvent intégralement en annexe du chapitre 9 (pages 51 - 59 des annexes).

Le Tableau 37 récapitule les mouvements formantiques dus à la labialisation, palatalisation et vélarisation dégagés à partir de la modélisation par VTDemo. Deux flèches signifient un mouvement plus important par rapport à la cible qu'une flèche.

	Labialisation			Palatalisation			Vélarisation		
	F1	F2	F3	F1	F2	F3	F1	F2	F3
u	Voyelle labiale			↓	↑	<div></div>	Voyelle vélaire - pharyngale		
ɑ	↓	↓	--	↓	↑↑				
i	--	↓	↓↓	Voyelle palatale			↑	↓	↓↓
y	Voyelle labiale						↑	↓	--

Tableau 37 : Les mouvements des formants F1, F2, (F3) durant la palatalisation de [u, ɑ], labialisation de [ɑ, i] et vélarisation de [i, y] par rapport à la cible. Deux flèches indiquent une variation formantique plus importante que celle des autres formants

Le Tableau 37 montre que les changements formantiques importants concernent les formants qui sont essentiellement dus à la cavité antérieure. Ainsi les variations les plus importantes concernent le F3 dans le cas de [i] (pour les stimuli où la cavité antérieure reste plus courte que la cavité postérieure) et le F2 dans le cas de [y] et de [ɑ] (pour les stimuli où la cavité antérieure reste plus longue que la cavité postérieure).

Les prédictions qui peuvent être dégagées à propos de la labialisation, palatalisation et vélarisation à partir de la modélisation articulatoire dans VTDemo sont les suivantes :

La <i>labialisation</i> des voyelles étirées engendre une baisse du premier formant pour les voyelles ayant un F1 élevé et des formants supérieurs qui sont dépendants de la cavité antérieure.
La <i>palatalisation</i> des voyelles postérieures engendre une baisse du premier formant et une élévation des formants supérieurs, associés à la cavité antérieure.
Enfin, la <i>vélarisation</i> des voyelles antérieures engendre une élévation du premier formant et une baisse des formants supérieurs, dus à la cavité antérieure.

Les prédictions dégagées à partir de VTDemo sont en accord avec la littérature (Brosnahan and Malmberg, 1970; Vaissière, 2009).

9.2 Coarticulation en français : Données réelles

Afin d'étudier les changements des valeurs formantiques dus au contexte phonétique à partir de nos données, nous avons d'abord calculé les moyennes formantiques de chaque voyelle (valeurs prises à la moitié de la durée vocalique, toute position d confondue). La construction du corpus et le calcul des moyennes sont récapitulés dans le Tableau 38. Le tableau avec les moyennes formantiques et les écarts types des voyelles prononcées en contextes [p, t, k, ʁ] se trouve à la page 60 et 61 des annexes.

Pour définir l'ampleur et le sens des mouvements formantiques des voyelles en contexte par rapport à ces mêmes voyelles prononcées en isolation, nous avons soustrait de chaque moyenne formantique des voyelles en contexte la valeur formantique prototypique, calculée à partir de

productions de voyelles isolées. Les différences de valeurs obtenues (renseignées dans le Tableau 39) nous permettent de chiffrer les mouvements formantiques dus au contexte consonantique symétrique : labial, dental, palato-vélaire et uvulaire. Elles sont alors soit positives, soit négatives selon le sens du mouvement (l'augmentation de la valeur du formant par rapport à la cible est exprimée par un chiffre positif et la baisse par un chiffre négatif). Les mouvements du formant F1 supérieur à 50 Hz et des formants F2, F3 et F4 supérieurs à 95 Hz par rapport à la cible sont grisés.

CORPUS																
10 voyelles orales du français [i, e, ε, y, ø, œ, u, o, ɔ, a], insérées dans des logatomes pVpVpVp, tVtVtVt, kVkVkVk, RVRVRVR d'une phrase cadre, exemple : « Le mot papapape peut bien coller. », répétées 4 fois																
LOCUTRICES																
10 Françaises non-méridionales natives ayant grandi ou longtemps séjourné en Ile-de-France																
CALCUL DE LA MOYENNE																
A partir de la valeur relevée à la moitié de la durée de chaque voyelle orale des logatomes trissyllabiques (10 locutrices* 4 répétitions* 3 syllabes)																

Tableau 38 : Corpus, locutrices et mesures des moyennes formantiques des voyelles du français prononcées en contextes p, t, k, R

	F1				F2				F3				F4			
	p	t	k	R	p	t	k	R	p	t	k	R	p	t	k	R
i	9	29	12	136	-153	-114	10	-30	-298	-366	-116	-508	-66	-92	-64	-196
y	14	22	18	114	-7	40	19	-341	82	210	27	281				
e	23	25	12	81	-202	-228	-77	-239	-324	-309	-223	-359				
ø	16	8	10	54	32	151	152	-255	32	137	-122	169				
ε	17	-21	-64	71	-152	-129	88	-244	-128	-56	-99	-96				
œ	18	-8	-37	62	47	183	240	-158	27	104	-203	59				
a	-6	-56	-97	15	278	515	683	133	-37	86	-140	-33				
u	25	16	20	58	82	440	76	-53								
o	15	6	21	21	100	393	153	19								
ɔ	57	29	31	77	170	389	207	50								

Tableau 39 : Distance acoustique (en Hertz) entre les valeurs moyennes formantiques des voyelles du français prononcées en contextes [p, t, k, ʁ] (calculées à partir de productions de 10 locutrices*4 répétitions*3 syllabes) et les valeurs des formants des mêmes voyelles prononcées en isolation (calculées à partir de productions de 10 locutrices*4 répétitions*3 valeurs). Les mouvements formantiques importants (supérieurs à 50 Hz pour F1 et à 95 Hz pour F2, F3 et F4) sont grisés

Dans l'étude des effets de coarticulation, nous émettons l'hypothèse que les valeurs formantiques des voyelles orales prononcées en contextes [p, t, k, ʁ] sont différentes de celles des voyelles prononcées en isolation.

L'hypothèse nulle consiste à dire qu'il n'existe pas de différences significatives entre les valeurs formantiques des voyelles prononcées en isolation et en contexte consonantique.

Pour mesurer l'effet des contextes nul (voyelles en isolation), labial, dental, palato-vélaire et uvulaire sur les formants F1, F2, (F3) des dix voyelles orales du français [i, e, ε, y, ø, œ, u, o, ɔ, a], nous avons conduit une analyse ANOVA à deux facteurs : voyelle et contexte. Nous avons observé l'effet du contexte (variable nominale à 5 modalités) et de la voyelle (variable nominale à 10 modalités) sur les valeurs formantiques ainsi que l'effet combiné des deux facteurs. Les résultats de l'analyse de variance sont donnés pour chaque formant séparément.

9.2.1 Coarticulation en français : Premier formant

La valeur du premier formant est essentiellement corrélée, comme exposé en 8.3.1, au degré de l'aperture, déterminée par la position verticale de la mâchoire et de la langue. Puisque la langue ou les lèvres empêche le passage de l'air durant l'occlusion des consonnes explosives [p, t, k] et que la mâchoire ou le corps de la langue ne peut s'abaisser qu'après le relâchement de l'occlusion, il n'est pas possible d'anticiper le geste d'ouverture avant le relâchement. En effet, une telle anticipation provoquerait le changement du mode d'articulation de la consonne. Ainsi l'assimilation de l'aperture – notamment dans le cas des voyelles ouvertes – est attendue.

Toutefois, l'assimilation de l'aperture semble moins ample que l'assimilation du lieu d'articulation: Stevens and House (1963) notent des changements de valeurs du premier formant - dus strictement à la coarticulation - moindres qu'au niveau du deuxième formant, du moins en anglais. Quand ces mouvements surviennent, les auteurs observent deux types de comportements : soit un abaissement du F1 dû à l'occlusion de la consonne (assimilation contextuelle), soit une centralisation acoustique des voyelles fermées visant la valeur de 500 Hz (qui est le F1 de la voyelle neutre, modélisée par un tube uniforme (Fant, 1960).

L'étude de Hillenbrand *et al.* (2001) sur les mêmes huit voyelles de l'anglais dans des syllabes CVC montre que la variation du premier formant due à l'effet consonantique est également minime mis à part le F1 des voyelles mi-ouvertes [ɛ] et [æ] qui baisse quand ces dernières se trouvent au contact de la consonne coronale [t] ou vélaire [k].

Le résultat est semblable pour le français : en considérant les effets des différents contextes phonétiques séparément (mais avec une seule condition de débit), Gendrot and Adda-Decker (2010) observent un mouvement essentiellement horizontal (changements de valeurs de F2), du moins pour [ɔ] et [a] (les auteurs n'exposent pas les résultats pour toutes les voyelles).

Notons enfin que les études qui s'intéressent à l'effet général des contextes consonantiques (en comparant les formants moyens de voyelles issues de tous les contextes confondus par rapport à ceux des voyelles en isolation) indiquent une centralisation claire (van Bergem, 1993). C'est par exemple le cas de l'étude des voyelles du suédois de Stalhammar *et al.* (1973) où l'effet du contexte et de l'accent est la centralisation des valeurs formantiques. Ainsi les voyelles produites en isolation sont à la périphérie alors que dès leur insertion dans des contextes consonantiques variés, elles tendent vers la voyelle neutre. La centralisation est la plus marquée dans le cas des voyelles qui apparaissent en syllabes inaccentuées.

Notre étude à partir de données réelles montre qu'en contexte labial, le premier formant de toutes les voyelles varie moins que les formants supérieurs (voir Tableau 39), ce qui est soutenu par la littérature et la modélisation par VTDemo. La différence entre le premier formant des voyelles en contexte symétrique labial pVp et celui des voyelles en isolation est ainsi inférieure à 50 Hz, mis à part

le [ɔ] mi-ouvert dont la valeur augmente de 57 Hz en moyenne par rapport à la cible. Ce résultat inattendu du contexte labial à augmenter le F1 d'une voyelle mi-ouverte pourrait s'expliquer par les règles orthopéiques qui déterminent la distribution semi-complémentaire des voyelles moyennes du français standard : les locutrices produisent le [ɔ] en isolation (ce qui n'est pas naturel) avec une valeur de F1 plus basse (corrélée essentiellement à un plus petit degré d'aperture) que lorsque la voyelle est insérée en contexte phonétique symétrique. La Figure 59 nous fait remarquer ce phénomène : Le F1 de [ɔ] en contextes p, t, k, R est en moyenne plus élevé que celui de [ɔ] en isolation (traits horizontaux). L'augmentation de la valeur du premier formant de [ɔ] en syllabe initiale – généralement produite avec une forte tension des articulateurs (Straka, 1964) – et en syllabe finale (allongée) des logatomes trisyllabiques par rapport à F1 de [ɔ] en isolation est systématique et concerne tous les environnements phonétiques étudiés.

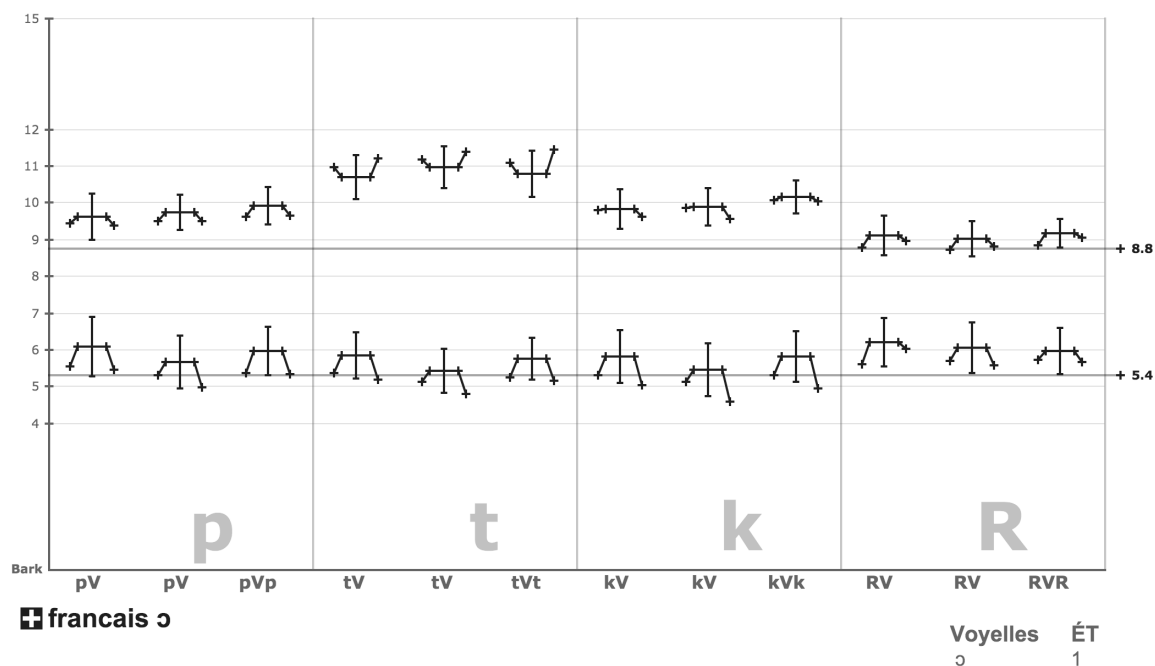


Figure 59 : Valeurs de F1 et F2 (en Bark) de [ɔ] produit en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par 10 Françaises natives et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1

En contexte dental tVt, le F1 de la seule voyelle [a] varie : il baisse – comme attendu- de 56 Hz.

En contexte palato-vélaire kVk, le changement de la valeur du premier formant supérieur à 50 Hz concerne les voyelles de grande aperture [ɛ, a] : le F1 de ces voyelles baisse de 64 Hz et 97 Hz respectivement. Ce résultat est en concordance avec l'étude de Hillenbrand *et al.* (2001).

Enfin, le contexte uvulaire RVR est le seul à provoquer des changements du premier formant importants et systématiques : ce dernier augmente de manière importante pour [i, y] (en moyenne de 136 Hz et 114 Hz respectivement), ce qui est visible sur le spectrogramme de la Figure 60, mais également pour [e, ɛ, ɔ] (où F1 augmente respectivement de 81 Hz, 71 Hz et 77 Hz) ainsi que pour les voyelles [ø, œ, u] (avec une élévation moyenne respectivement de 54 Hz, 62 Hz, 58 Hz).

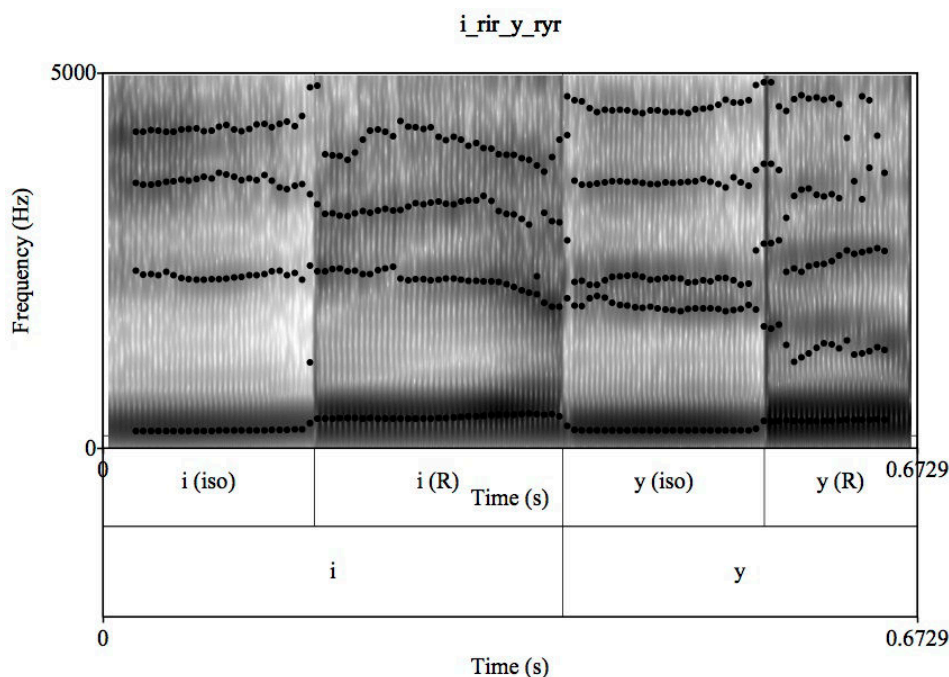


Figure 60 : Spectrogramme des voyelles [i, y] du français prononcées isolément et en contexte RVR. Notons que le contexte uvulaire engendre une élévation du F1 et un abaissement des formants principalement dus à la cavité antérieure (F3 pour [i] et F2 pour [y]).

L'élévation de F1 en seul contexte uvulaire correspond également à la littérature car selon les lois acoustiques, toute constriction dans la moitié antérieure du conduit vocal (donc articulation labiale, dentale ou palato-vélaire) conduit vers une baisse du premier formant et seulement les articulations pharyngales ou glottales engendrent une élévation du F1 (Fant, 1960; Stevens, 1998).

En revanche, le premier formant des voyelles [a, o] n'est pas affecté de manière importante par le contexte uvulaire (leur F1 n'augmente que de 15 Hz et 21 Hz respectivement).

Le graphe de la Figure 61 illustre les variations de F1 dues au contexte pour chaque voyelle séparément.

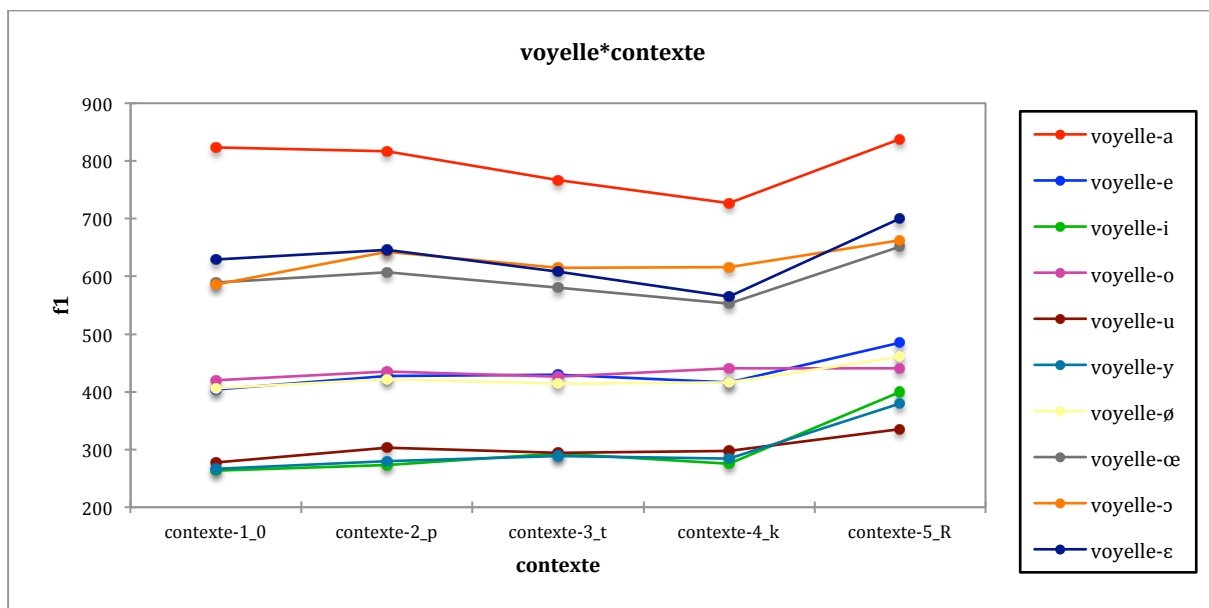


Figure 61 : Le F1 des dix voyelles orales du français [i, e, ε, y, ø, œ, u, o, ɔ, a] prononcées isolément (0) et en contextes [p, t, k, ʁ] dans des logatomes CVCVCVC par 10 locutrices*4 répétitions.

Le graphe de la Figure 61 montre l'évolution du premier formant des dix voyelles orales du français actuel [i, e, ε, y, ø, œ, u, o, ɔ, a] selon si la voyelle avait été prononcée isolément ou en contextes labial, palatal, palato-vélaire ou uvulaire. Le contexte nul (voyelle en isolation) sert de référence. Nous remarquons que le F1 de la voyelle [a] et [ε] en contexte kVk est plus bas que celui de [a] et [ε] isolés. En revanche, le F1 de toutes les voyelles, et notamment celui de [ε, ɔ, œ, e, y, i] est plus élevé en contexte uvulaire. Les autres contextes phonétiques engendrent des mouvements formantiques moindres. Ces résultats sont en concordance avec la littérature mentionnée ci-dessus.

Le résultat de l'analyse de variance donné par la valeur de l'ANOVA $F(49, 5142) = 761$ avec $p < 0,05$ du Tableau 40 montre un effet global du contexte consonantique sur les valeurs du premier formant des dix voyelles orales du français.

F1	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	49	137422738,921	2804545,692	761,215	< 0,0001
Erreur	5142	18944684,169	3684,303		

Tableau 40 : Résultat de l'ANOVA à 2 facteurs qui indique l'effet du contexte sur le formant F1 des dix voyelles orales du français (10 locutrices*4 répétitions*3 syllabes)

Le Tableau 41 montre un effet significatif du premier facteur – Voyelle – sur les valeurs du F1 avec $F(9, 5142) = 3962$, $p < 0,05$. Il montre également un effet du deuxième facteur – Contexte – sur les valeurs du F1, avec $F(4, 5142) = 283$, $p < 0,05$. Finalement, il existe une interaction entre les deux facteurs – Voyelle*Contexte – avec $F(36, 5142) = 14$, $p < 0,05$.

F1	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle	9	131360623,958	14595624,884	3961,571	< 0,0001
contexte	4	4176240,315	1044060,079	283,381	< 0,0001
voyelle*contexte	36	1885874,647	52385,407	14,219	< 0,0001

Tableau 41 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Contexte et Voyelle*Contexte sur le F1 des voyelles orales du français (10 locutrices*4 répétitions*3 syllabes)

Pour vérifier à quels groupes est dû l'effet global combiné des deux facteurs Voyelle*Contexte, nous avons procédé au test a posteriori de Fisher. Le résultat complet se trouve dans les annexes (page 62) et il indique que les valeurs moyennes formantiques des voyelles en isolation et celles des voyelles en contexte sont significativement différentes au niveau du premier formant :

- en contexte labial, pour les voyelles [e, u, ɔ] (contexte-0 vs contexte-p)
- en contexte dental, pour les voyelles [i, e, y, ɔ, a] (contexte-0 vs contexte-t)
- en contexte palato-vélaire, pour les voyelles [ε, œ, ɔ, a] (contexte-0 vs contexte-k)
- en contexte uvulaire, pour toutes les voyelles, mis à part le [o, a] (contexte-0 vs contexte-R)

Ainsi, certains contextes phonétiques provoquent de légers mouvements de F1 de certaines voyelles par rapport à leurs cibles.

9.2.2 Coarticulation en français : Deuxième formant

Dans le cas des voyelles dont le deuxième formant dépend essentiellement de la cavité antérieure, comme [y] et [a] ou les voyelles postérieures, les variations de F2 dépendent de la nature des sons successifs : les mouvements par rapport à la cible sont moins amples lorsque la voyelle

partage avec la consonne environnante le même lieu de constriction et arrondissement des lèvres. En revanche, dès qu'une voyelle est entourée de consonnes articulées à un endroit du conduit vocal éloigné (comme par exemple la voyelle postérieure [u] en contexte dental) ou avec une configuration des lèvres différente (comme la voyelle antérieure arrondie [y] en contexte autre que labial), un plus grand mouvement formantique est attendu. Il est cependant important de souligner que la position du dos de la langue et des lèvres peut être anticipée durant la consonne, contrairement au degré d'aperture (Flemming, 2006; Hirsch *et al.*, 2004). Ainsi lors de la production des consonnes labiales, la langue, qui ne participe pas à l'occlusion de la consonne, peut adopter des formes diverses, dont celle requise pour la voyelle. Le point d'articulation des consonnes vélaires s'adapte également largement à celui de la voyelle adjacente (pour cette raison, l'occlusive [k] est dénommée « consonne palato-vélaire »). Enfin, l'anticipation de la position de la langue pour la voyelle durant la production d'une dentale qui la précède est plus limitée mais on sait depuis Öhman (1966) qu'elle est possible.

Bothorel *et al.* (1986) ont montré par la cinéradiographie des voyelles et consonnes du français que l'assimilation régressive (autrement dit coarticulation anticipante) est la plus productive en français. Ils ont montré que les consonnes peuvent être par anticipation palatalisées (par exemple [kʲ] dans [ekʲi]), vélarisées ([ɫ] dans [ɫwi]) ou labialisées ([dʷ] dans [edʷy]). Ils ont également reporté des changements dans l'articulation des voyelles en fonction des consonnes qui les entourent : ces dernières modifient leur lieu d'articulation (par exemple [ɔ] est antériorisé lorsqu'il se trouve entouré de consonnes dentales comme dans [dɔn]) ou leur aperture (par exemple [a] de [vɑ] est produit avec une aperture plus petite que [a] final de [aba]). Selon Vaissière (2007), l'anticipation de voyelles (V) durant des consonnes (C) dans des séquences CV en français est tellement forte que toutes les consonnes sont phonétiquement palatalisées, vélarisées, pharyngalisées ou labialisées en fonction de la voyelle qui les suit.

Les effets de consonnes sur les voyelles environnantes ont été étudiés pour l'anglais par Stevens and House (1963) et plus récemment par Hillenbrand *et al.* (2001). Selon les auteurs, le F2 des voyelles antérieures de l'anglais américain baisse par rapport à la cible lorsque ces dernières sont prononcées en contextes labial et alvéolaire. En revanche, celui des voyelles postérieures augmente en contexte alvéolaire et dans une moindre mesure en contexte vélaire. Un tableau comparant l'ampleur des variations dues à l'effet consonantique reportées par les deux auteurs se trouve en annexe du chapitre 9 (pages 61 – 62 des annexes).

La coarticulation en français a été étudiée par Durand (1985). En se basant sur la lecture de phrases comportant de vrais mots de même que des non-mots par deux locuteurs masculins, ce dernier s'intéressait à l'effet des consonnes occlusives [p, t, k, b, d, g] sur la fréquence centrale des formants vocaliques. Le résultat montre que le deuxième formant de toutes les voyelles en contexte bilabial baisse, mis à part celui du [i], qui est dû à la cavité postérieure. En contexte dental, les voyelles ayant un F2 supérieur ou égal à 2000 Hz enregistrent une baisse de ce dernier et au contraire, celles ayant un F2 inférieur à 1800 Hz indiquent son élévation. En contexte palato-vélaire, l'auteur constate une élévation du F2 des voyelles ayant un deuxième formant inférieur à 2000 Hz et note que [y] est la seule voyelle dont le F2 ne change pas au contact de [k] ou [g].

Les plus grands déplacements du F2 par rapport à la cible reportés par les études se produisent en contexte coronal. Une forte réduction vocalique due au contexte coronal est également observée en français par Gendrot and Adda-Decker (2010). Notons qu'en français et en tchèque, où l'articulation de [t] est essentiellement dentale [t̪] (Dart, 1998; Skarnitzl, 2014) contrairement à l'articulation alvéolaire [t̬] de l'anglais américain, le F2 vocalique se dirige en présence de cette dernière vers « le locus dental » - dénomination utilisée par Delattre (1963) - qui se trouve aux alentours de 1800 Hz.

Nos résultats du Tableau 39 de la page 156 montrent que les variations du deuxième formant dues au contexte phonétique immédiat sont plus grandes que celles du premier formant, ce qui est un résultat attendu.

Le graphe de la Figure 62 illustre les variations du deuxième formant des dix voyelles orales du français [i, e, ε, y, ø, œ, u, o, ɔ, a] en fonction du contexte dans lequel la voyelle avait été prononcée. Le contexte nul (0) correspond aux voyelles prononcées isolément et il sert de référence.

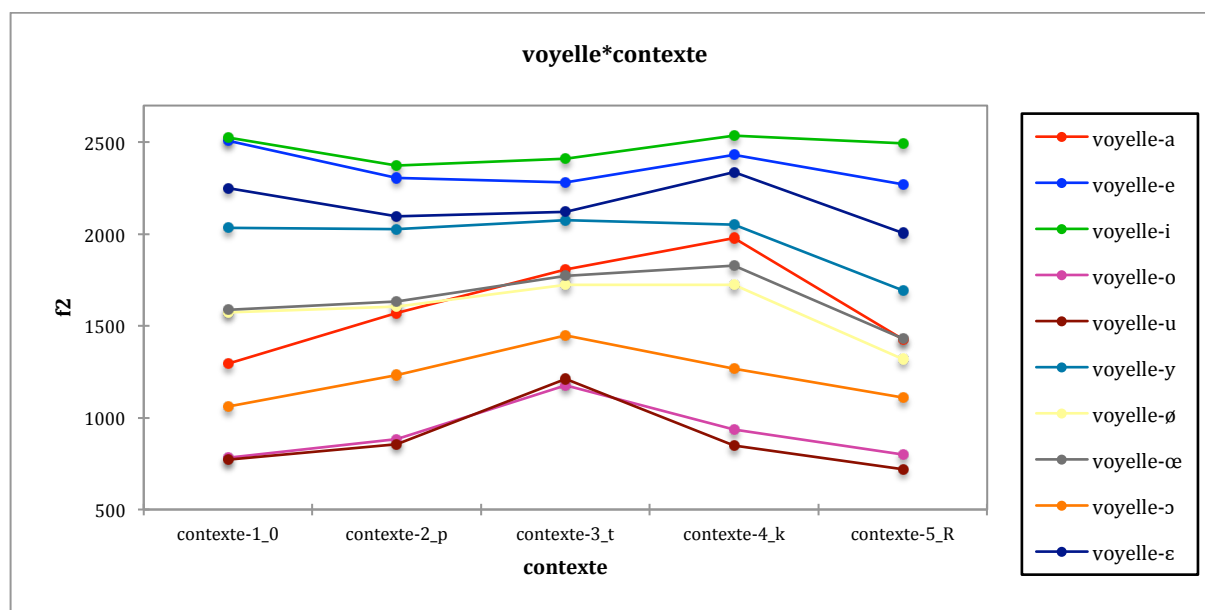


Figure 62 : Le F2 des dix voyelles orales du français [i, e, ε, y, ø, œ, u, o, ɔ, a] prononcées isolément (0) et en contextes [p, t, k, ʁ] dans des logatomes CVCVCVC (10 locutrices*4 répétitions)

La Figure 62 nous fait remarquer que le contexte uvulaire provoque en général par rapport à la cible une baisse du deuxième formant, associé à la cavité antérieure qui se prolonge, (mis à part le [a]), le contexte palato-vélaire est responsable d'une élévation du deuxième formant associé à la cavité antérieure, notamment celle du [a], et les contextes labial et dental engendrent une baisse du deuxième formant des voyelles antérieures étirées et une élévation importante du deuxième formant des voyelles postérieures.

Le résultat de l'analyse de variance indiquant l'effet global significatif du contexte phonétique sur la valeur du deuxième formant des dix voyelles orales du français est exprimé par la valeur ANOVA $F(49, 5142) = 1440$ et $p < 0,05$ (voir le Tableau 42).

Source	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	49	1512730873,921	30872058,651	1440,314	< 0,0001
Erreur	5142	110214974,578	21434,262		

Tableau 42 : Résultat de l'ANOVA à 2 facteurs qui indique l'effet du contexte sur le formant F2 des dix voyelles orales du français (10 locutrices*4 répétitions*3 syllabes)

Le Tableau 43 montre un effet significatif du facteur « Voyelle » sur la valeur du F2 avec $F(9, 5142) = 7255$, $p < 0,05$, de même que du facteur « Contexte » avec la valeur ANOVA $F(4, 5142) = 731$, $p < 0,05$, et également des deux facteurs combinés « Voyelle*Contexte » avec $F(36, 5142) = 65$, $p < 0,05$.

Source	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle	9	1399609920,276	155512213,364	7255,310	< 0,0001
contexte	4	62676733,574	15669183,393	731,034	< 0,0001
voyelle*contexte	36	50444220,072	1401228,335	65,373	< 0,0001

Tableau 43 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Contexte et Voyelle*Contexte sur le F2 des voyelles du français (10 locutrices*4 répétitions*3 syllabes)

Le test a posteriori de Fisher indique à quels groupes est due l'interaction des facteurs « Voyelle*Contexte ». Le résultat complet se trouve dans les annexes à la page 63 et il montre que les valeurs de F2 significativement différentes de celles de voyelles en isolation concernent :

- en contexte labial, toutes les voyelles, mis à part le [y, ø, œ] (contexte-0 vs contexte-p)
- en contexte dental, toutes les voyelles, mis à part le [y] (contexte-0 vs contexte-t)
- en contexte palato-vélaire, toutes les voyelles, mis à part le [i, y] (contexte-0 vs contexte-k)
- en contexte uvulaire, toutes les voyelles, mis à part le [i, o, ɔ] (contexte-0 vs contexte-R)

Ainsi les changements de valeurs du deuxième formant sont minimales dans des suites combinant la consonne labiale [p] avec les voyelles antérieures labiales ([pyp], [pøp], [pœp]), ou la consonne palato-vélaire [k] avec les voyelles antérieures fermées comme dans [kik] et [kyk] ce que nous pouvons observer dans le spectrogramme de la Figure 63. De même, la consonne uvulaire [ʁ] ne modifie pas significativement le deuxième formant des voyelles postérieures [o, ɔ] ni de [i] dont le F2 dépend essentiellement de la cavité postérieure. Enfin, la consonne dentale [t] conserve un deuxième formant du [y] proche de la cible.

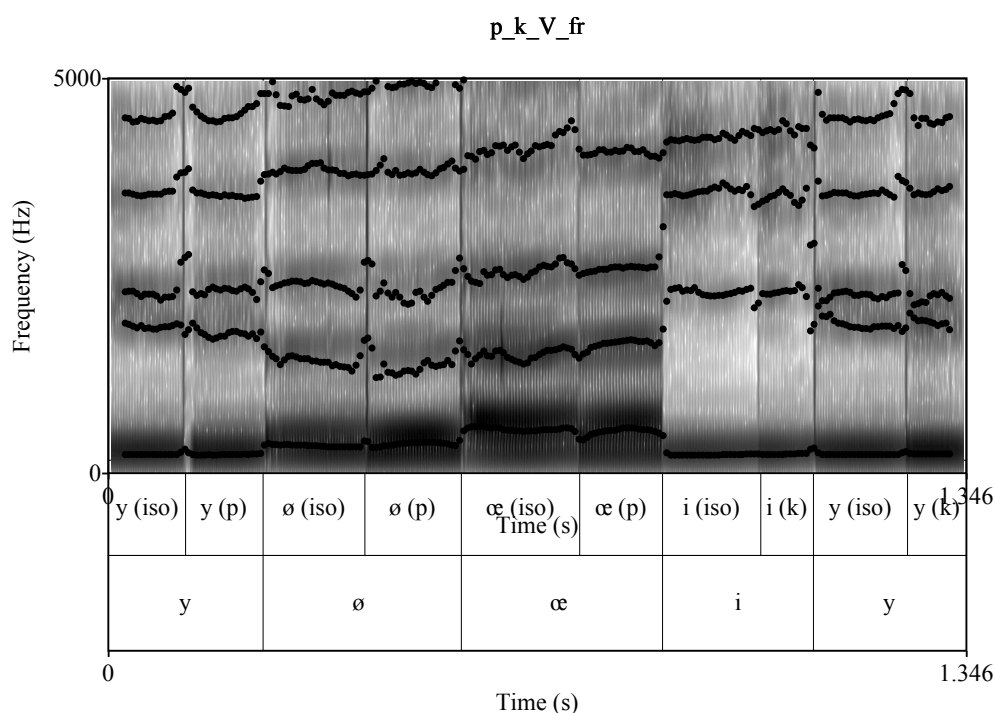


Figure 63 : Spectrogramme des voyelles [y, ø, œ, i] du français prononcées par une Française isolément et en contextes pVp et/ou kVk qui provoquent par rapport à la cible des mouvements minimes, comme expliqué dans le paragraphe ci-dessus et indiqué dans le Tableau 39

Les différents effets des contextes labial, dental, palato-vélaire et uvulaire sur le F2 des voyelles orales du français sont élucidés ci-dessous.

9.2.2.1 Effet du contexte labial sur le F2 (en français)

Le triangle F1/F2 (en Bark) montre par rapport à la cible (en trait plein) les mouvements formantiques des voyelles du français induits par le contexte symétrique labial [p] (en trait pointillés). L'écart type affiché est de un.

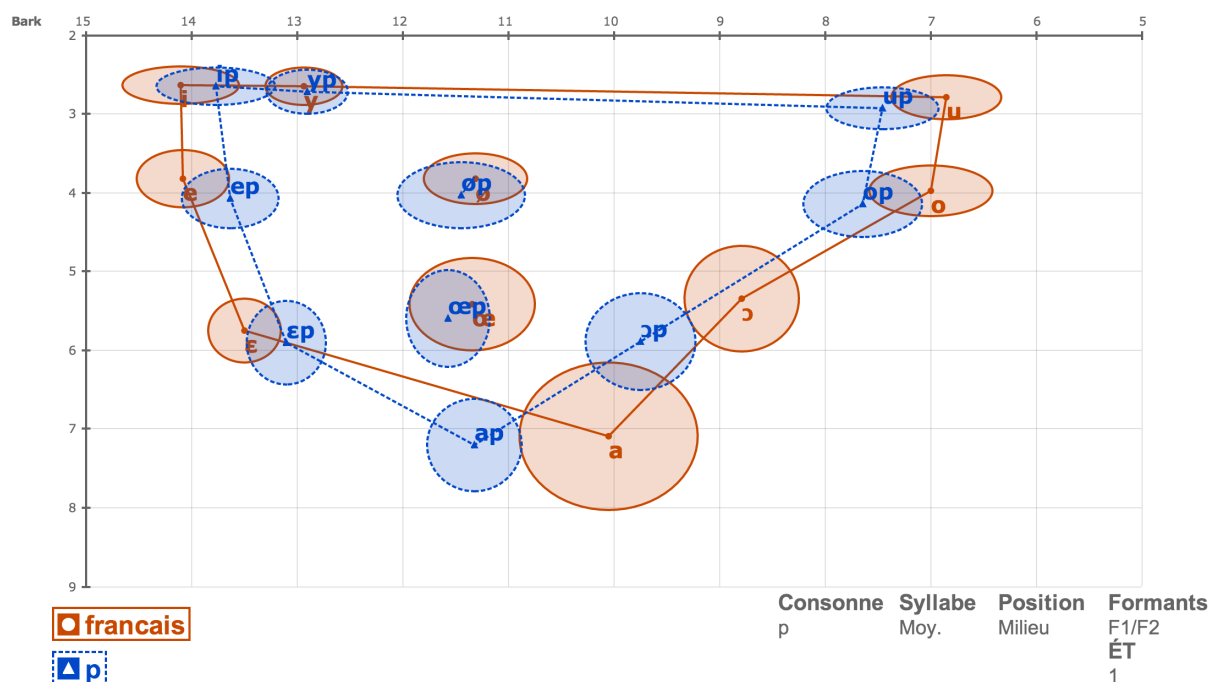


Figure 64 : Triangle vocalique des voyelles orales du français en isolation (trait plein) et en contexte symétrique labial (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 10 Françaises * 4 répétitions de chaque voyelle * valeur prise à la moitié de la durée vocalique

La Figure 64 montre que le contexte labial provoque un abaissement du deuxième formant des voyelles antérieures étirées par rapport à la cible : cette baisse du F2 est en moyenne de 153 Hz pour la voyelle [i], de 202 Hz pour [e] et de 152 Hz pour [ε]. En revanche, les voyelles antérieures arrondies sont marquées par des variations du F2 négligeables car elles partagent avec la consonne labiale [p] le même degré d'arrondissement aux lèvres. Le contexte labial engendre une augmentation progressive du deuxième formant des voyelles postérieures selon le degré d'aperture (et donc le degré d'arrondissement et protrusion), ce qui est un résultat inattendu. Dans le cas du [u] fermé et [o] mi-fermé, la montée du F2 ne dépasse pas 100 Hz. Ce résultat est attendu car les articulateurs bougent peu (les lèvres restent arrondies et la langue est libre à prendre la position de la voyelle). En revanche, dans le cas de la voyelle [ɔ], moins arrondie que [u] et [o], et [a], non arrondie, le F2, dû à la cavité antérieure qui se raccourcit par l'avancement du dos de la langue (au lieu de s'allonger par l'action des lèvres comme prévu par la modélisation), augmente en moyenne de 170 Hz et 278 Hz respectivement. Nous pouvons observer cette élévation importante du deuxième formant dans le spectrogramme de la Figure 65.

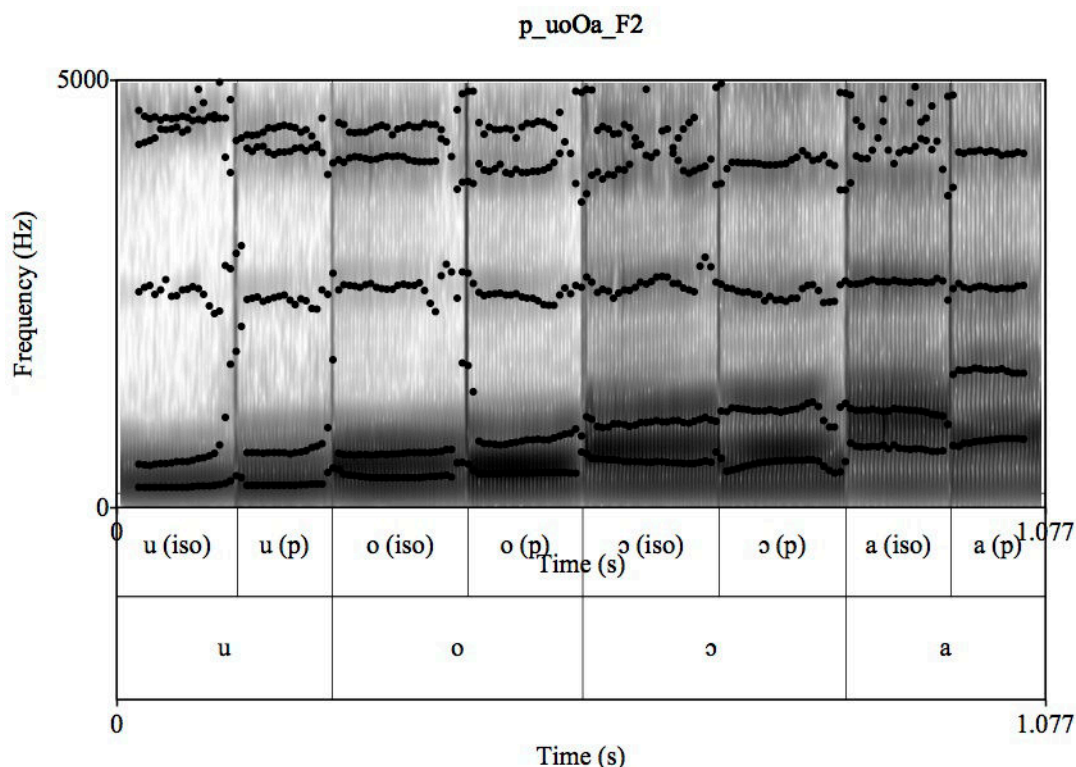


Figure 65 : Spectrogramme des voyelles [u, o, ɔ, a] du français prononcées isolément et en contexte pVp par une Française native. Notons que le contexte labial engendre une augmentation du F2 des voyelles postérieures de même que du [a] qui est pharyngal en isolation

La Figure 65 est un spectrogramme des voyelles [u, o, ɔ, a] prononcées isolément (iso) et en contexte labial (p). Il est à noter que le contexte symétrique pVp provoque une élévation du deuxième formant des trois voyelles postérieures et en particulier celui du [a]. Ce résultat n'est pas en accord avec la modélisation de la labialisation du [a] dans VTDemo qui prévoit une baisse du deuxième formant due à l'arrondissement des lèvres mais s'explique par la réalisation pharyngale de [a] en isolation (voir les cours de phonétique expérimentale de Vaissière).

9.2.2.2 Effet du contexte dental sur le F2 (en français)

Le triangle F1/F2 (en Bark) montre par rapport à la cible (en trait plein) les mouvements formantiques des voyelles du français induits par le contexte symétrique dental [t] (en trait pointillés). L'écart type affiché est de un.

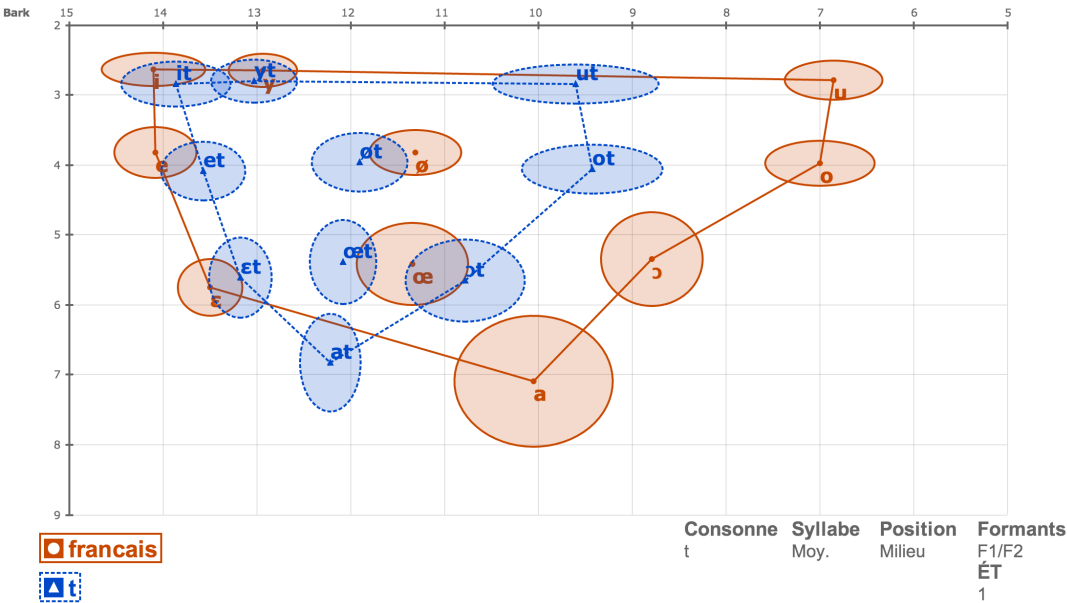


Figure 66 : Triangle vocalique des voyelles orales du français en isolation (trait plein) et en contexte symétrique dental (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 10 Français *4 répétitions de chaque voyelle*valeur prise à la moitié de la durée vocalique

Le contexte dental a un effet important sur le deuxième formant dû à la cavité antérieure, comme l'indique le Tableau 39 de la page 156 et la Figure 66 : toutes les voyelles sont en effet soumises à de grandes variations du F2 (qui sont néanmoins moins importantes que celles du F3 dans le cas des voyelles antérieures [i, e]), mis à part le [y] dont la valeur n'augmente que de 40 Hz. Alors que la valeur de F2 des voyelles antérieures étirées baisse, celle des voyelles antérieures labiales ainsi que des voyelles postérieures, dont la cavité antérieure se raccourcit, augmente en visant le locus dental de 1800 Hz.

L'élévation du F2 est maximale au niveau des voyelles postérieures [u, o, ɔ] et du [a], où la différence par rapport à la cible est respectivement de 440 Hz, 393 Hz, 389 Hz et 515 Hz en moyenne.

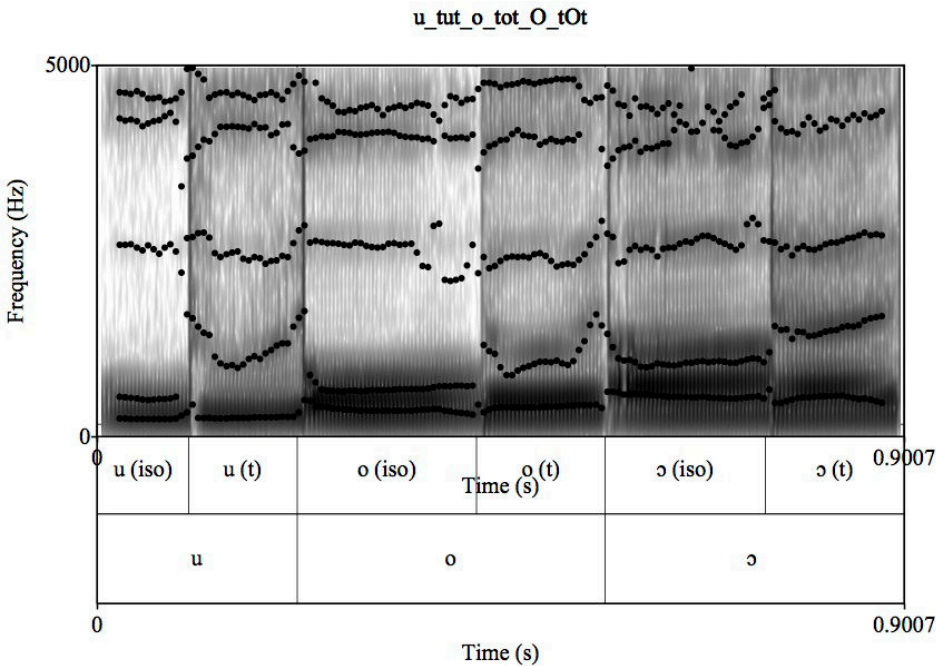


Figure 67 : Spectrogramme des voyelles [u, o, ɔ] du français prononcées isolément et en contexte tVt. Notons que le contexte labial engendre une augmentation importante du F2 des voyelles postérieures du français. Notons que le caractère focal des voyelles postérieures, qui rapprochent les formants F2-F1 en contexte nul, est particulièrement affaibli en contexte [t]

Le spectrogramme de la Figure 67 montre l'évolution des formants des voyelles [u, o, ɔ] prononcées par une Française native en isolation (iso) et en contexte symétrique dental (t). Nous y remarquons une élévation importante du deuxième formant vocalique due au raccourcissement de la cavité antérieure de laquelle il dépend.

9.2.2.3 Effet du contexte palato-vélaire sur le F2 (en français)

Le triangle F1/F2 (en Bark) montre par rapport à la cible (en trait plein) les mouvements formantiques des voyelles du français induits par le contexte symétrique palato-vélaire [k] (en trait pointillés). L'écart type affiché est de un.

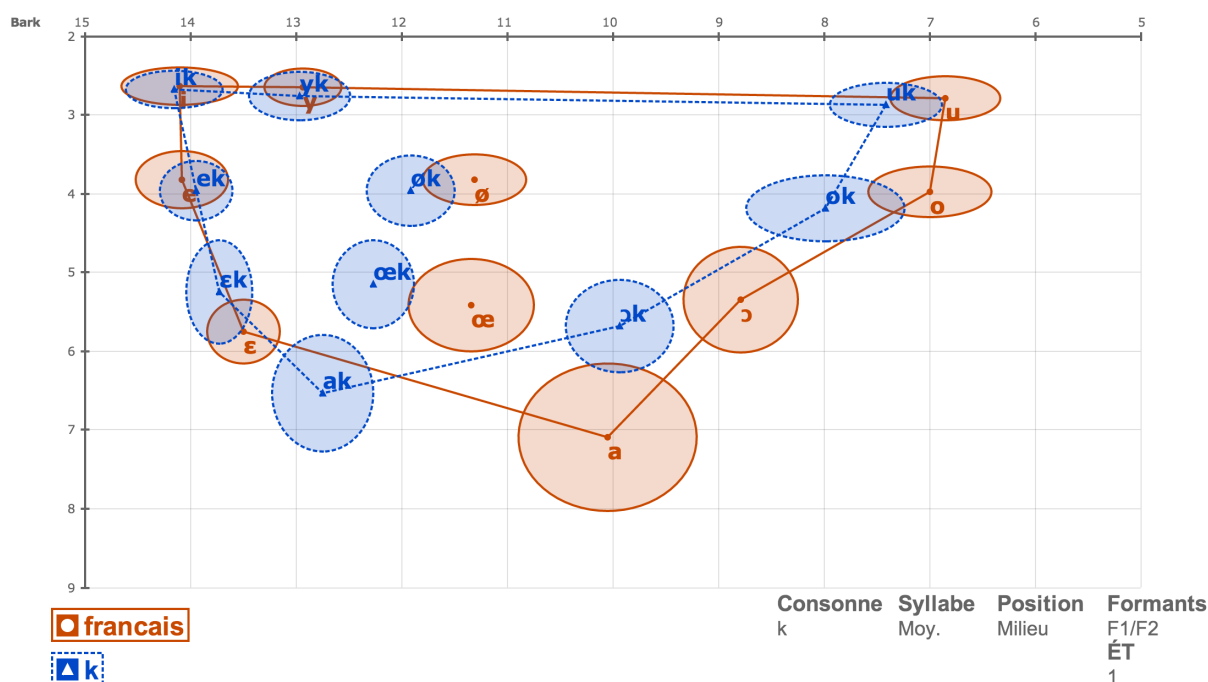


Figure 68 : Triangle vocalique des voyelles orales du français en isolation (trait plein) et en contexte symétrique palato-vélaire (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 10 Françaises *4 répétitions de chaque voyelle* valeur prise à la moitié de la durée vocalique

Le contexte palato-vélaire augmente systématiquement la valeur du deuxième formant de toutes les voyelles (voir Tableau 39 de la page 156 et Figure 68), à l'exception du [e] dont la valeur baisse de 77 Hz en moyenne (mouvement formantique cependant peu important). Une montée particulièrement élevée du deuxième formant est remarquée pour la voyelle [a] : le F2 de cette dernière dépend essentiellement de la cavité antérieure, raccourcie en contact avec la consonne palato-vélaire [k]. L'élévation moyenne est de 683 Hz par rapport au F2 de [a] isolé. Ainsi alors que la voyelle [a] prononcée isolément est postérieure, elle est centrale en parole continue.

9.2.2.4 Effet du contexte uvulaire sur le F2 (en français)

Le triangle F1/F2 (en Bark) montre par rapport à la cible (trait plein) les mouvements formantiques des voyelles du français induits par le contexte symétrique uvulaire [ʁ] (trait pointillé). L'écart type affiché est de un.

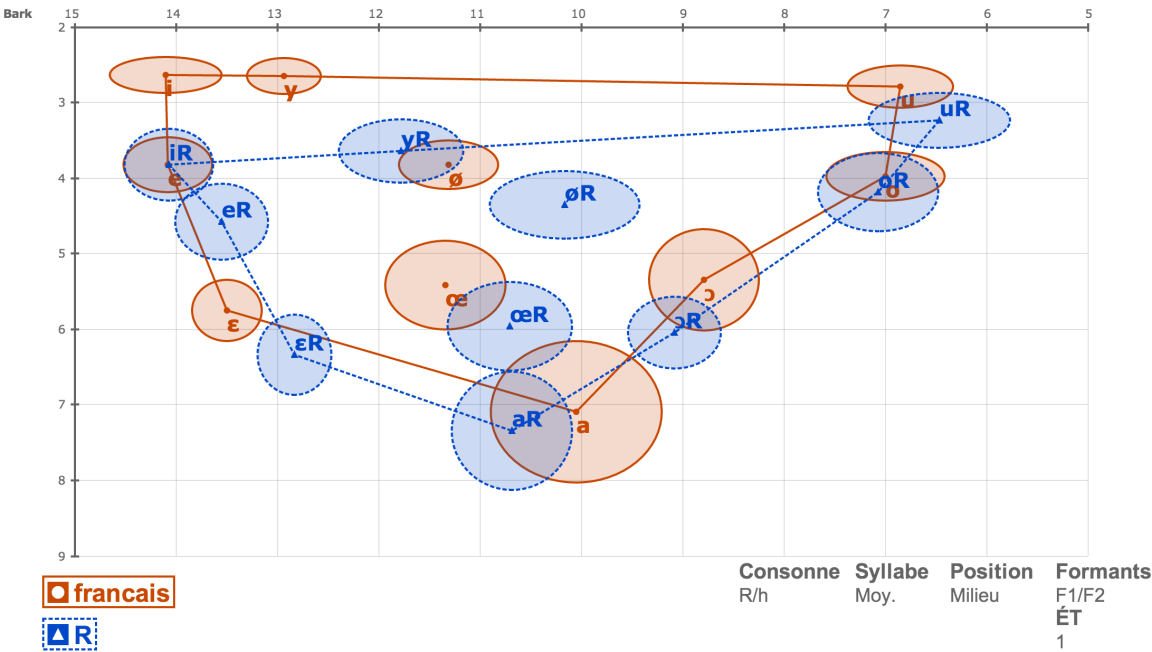


Figure 69 : Triangle vocalique des voyelles orales du français en isolation (trait plein) et en contexte symétrique uvulaire (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 10 Françaises *4 répétitions de chaque voyelle*valeur prise à la moitié de la durée vocalique

Enfin, comme l’indique la Figure 69, le contexte uvulaire fait baisser la valeur du formant dû à la cavité antérieure des voyelles antérieures. Le deuxième formant de toutes les voyelles antérieures, à l’exception du [i], est ainsi affecté : on note une baisse par rapport à la cible, illustré dans le spectrogramme de la Figure 70 et indiqué dans le Tableau 39 de la page 156, qui est en moyenne de 341 Hz, 239 Hz, 255 Hz et 244 Hz dans le cas des voyelles [y, e, ø, ε] respectivement, puis de 158 Hz dans le cas de [œ] et enfin seulement de 30 Hz pour [i]. Le mouvement négligeant dans le cas du deuxième formant de [i] est attendu car ce dernier dépend de la cavité postérieure, non affectée par la coarticulation.

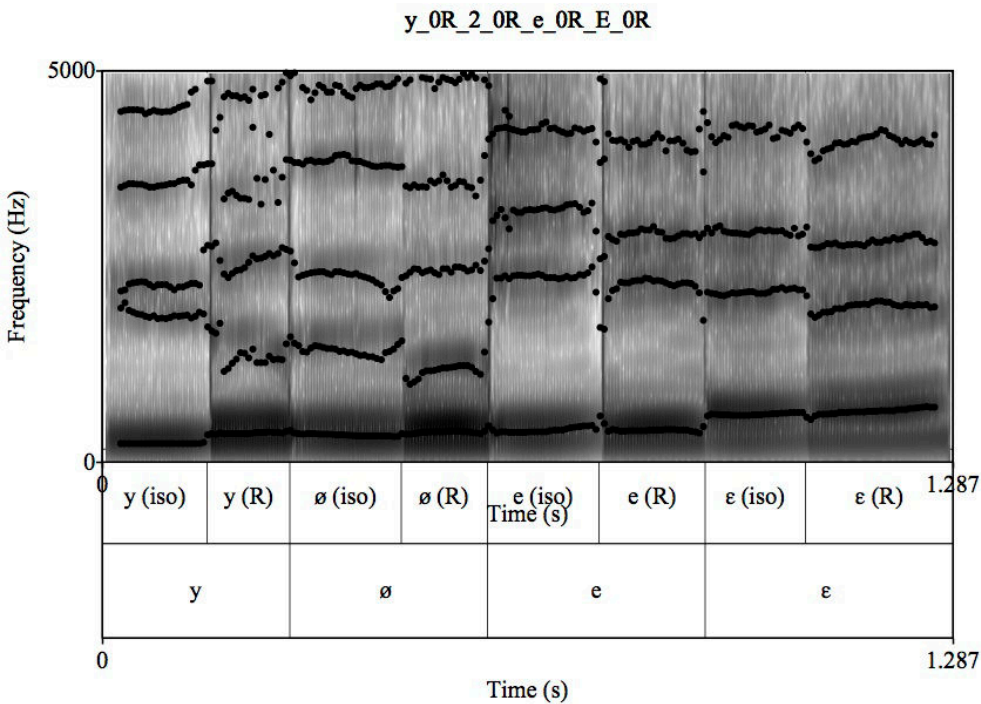
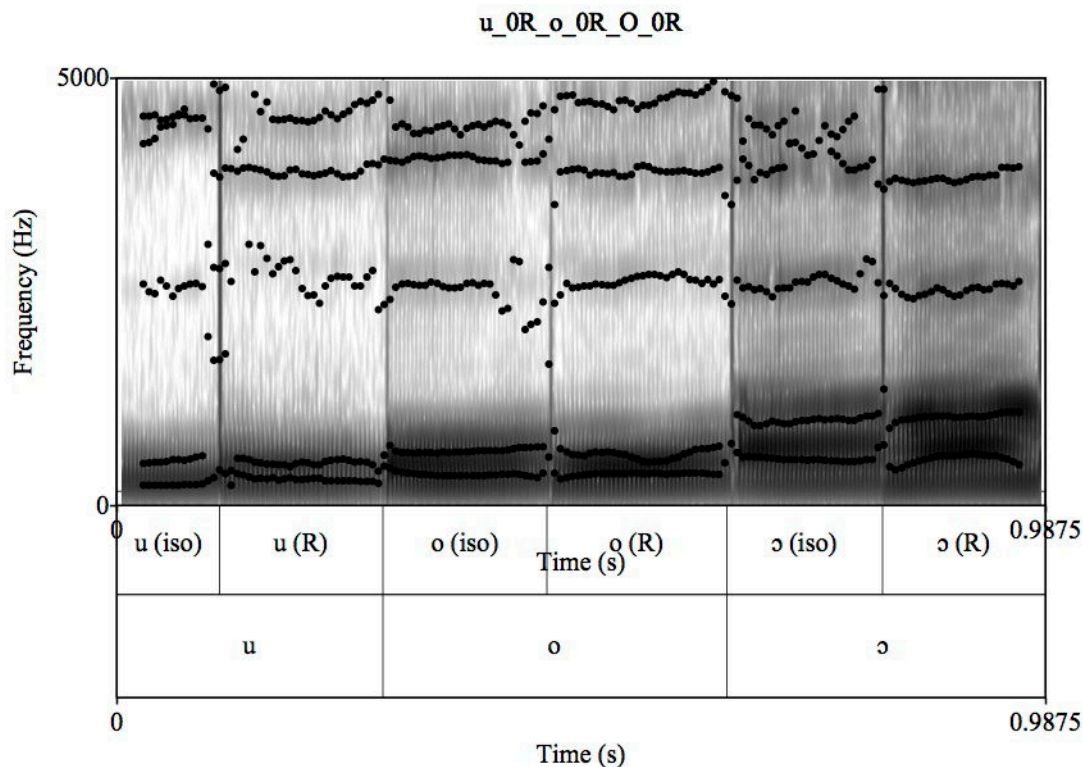


Figure 70 : Spectrogramme des voyelles [y, ø, e, ε] du français prononcées par une Française native isolément et en contexte RVR qui engendre une baisse du F2

La Figure 70 est un spectrogramme des voyelles [y, ø, e, ε] du français prononcées par une Française native isolément (iso) et en contexte symétrique uvulaire (R). Nous observons que le deuxième formant baisse de manière plus importante dans le cas des voyelles labiales [y] et [ø] qui l'associent principalement à la cavité antérieure que dans le cas des voyelles étirées [e, ε] où le F2 n'est affilié à la cavité antérieure qu'en partie. Notons également que la coarticulation des voyelles antérieures labiales avec une consonne postérieure requiert plusieurs gestes articulatoires qui concernent les lèvres (qui s'étirent) et la langue (qui recule). A l'inverse, dans le cas d'une coarticulation des voyelles antérieures non labiales avec une consonne uvulaire, les lèvres ne bougent pas.

En ce qui concerne les voyelles postérieures [u, o, ɔ], comme visibles sur le spectrogramme de la Figure 71, la consonne uvulaire provoque des changements de la valeur du deuxième formant des voyelles postérieures minimales (baisse de F2 par rapport à la cible inférieure à 100 Hz). Ceci est attendu car les deux sons en contact partagent le même lieu d'articulation et les articulateurs bougent minimalement durant la coarticulation de l'uvulaire [ʁ] avec les voyelles postérieures. Notons néanmoins le sens de ces petits déplacements formantiques : alors que le F2 de [u] baisse, ce qui est attendu, celui du [o] et notamment du [ɔ] augmente en moyenne de quelques dizaines de Hertz. Ces déplacements ne s'expliquent plus par l'assimilation contextuelle mais plutôt par la centralisation acoustique.



Ainsi, le F2 des voyelles postérieures de moyenne ou grande aperture - [ɔ] et notamment [a] (réalisée pharyngale en isolation) – s’élève lorsque ces dernières apparaissent en contexte et cela même en contextes symétriques labial et uvulaire. De plus, le F1 de [ɔ] inséré dans les quatre contextes est systématiquement plus élevé que celui de [ɔ] en isolation. Ce comportement inattendu des voyelles [a] et [ɔ] isolées nous amène à les considérer comme ayant une identité acoustique particulière et indique que nous ne pouvons pas nous en servir comme référence dans l’étude de la coarticulation. Cette observation a déjà été faite pour le /a/ par Vaissière dans ses cours de phonétique expérimentale de Master. Le comportement du [ɔ] en contexte dont le patron formantique ne peut s’expliquer ni par l’assimilation contextuelle ni par la centralisation acoustique est probablement relié à l’enfreint des règles orthoépiques du français standard selon lesquelles cette voyelle n’apparaît jamais en syllabe finale ouverte. Ainsi, le comportement inattendu du [ɔ] s’explique plutôt par le corpus comportant des voyelles isolées.

9.2.3 Coarticulation en français : Troisième formant

Le troisième formant n’a été étudié que pour les voyelles antérieures et la voyelle [a]. Les voyelles particulièrement sensibles à l’influence du contexte consonantique au niveau de leur troisième formant sont le [i] et le [e], comme nous pouvons le lire dans le Tableau 39 de la page 156, ce qui est attendu car le F3 est essentiellement une résonance de la cavité antérieure pour ces deux voyelles.

Le graphe de la Figure 72 illustre les variations du troisième formant dues au contexte pour chaque voyelle séparément.

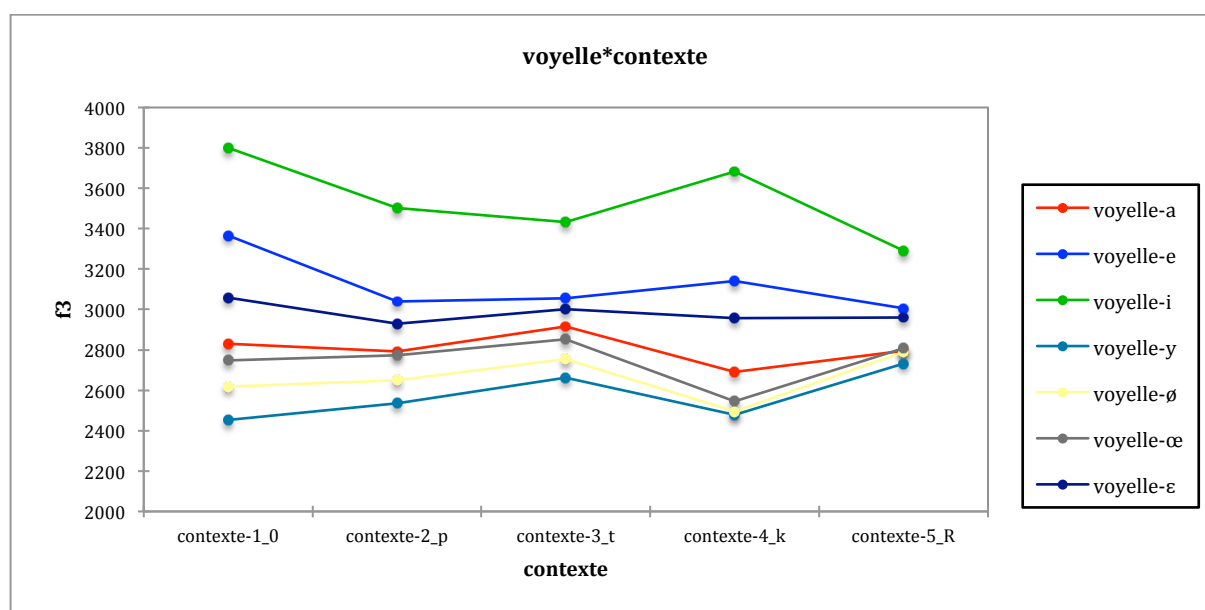


Figure 72 : Le F3 moyen des voyelles orales du français [i, e, ε, y, ø, œ, a] prononcées isolément (0) et en contextes [p, t, k, ɾ] dans des logatomes CVCVCVC (10 locutrices* 4 répétitions* valeur prise à la moitié de la durée vocalique)

Le graphe de la Figure 72 montre les variations du troisième formant des voyelles [i, e, ε, y, ø, œ, a] selon le contexte dans lequel elles avaient été prononcées (nul 0, labial p, dental t, palato-vélaire k ou uvulaire R). Alors que le contexte uvulaire provoque une baisse du F3 des voyelles antérieures étirées et une augmentation du F3 des voyelles antérieures labiales, le contexte palato-vélaire provoque une baisse de F3 de toutes les voyelles étudiées. Le contexte labial et dental engendre une baisse de F3 du [i] et [e] qui est essentiellement affilié à la cavité antérieure.

Le résultat du test ANOVA montre que l'effet global du contexte phonétique sur le troisième formant des dix voyelles est également significatif, comme nous pouvons le voir dans le Tableau 44, car $F(49, 5142) = 207, p < 0,05$.

Source	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	49	381926879,028	7794426,103	207,467	< 0,0001
Erreur	5142	193181870,957	37569,403		

Tableau 44 : Résultat de l'ANOVA à 2 facteurs qui mesure l'effet du contexte sur le formant F3 des dix voyelles orales du français (10 locutrices*4 répétitions*3 syllabes)

La valeur ANOVA indiquant l'effet du premier facteur « Voyelle » sur le F3 est $F(9, 5142) = 944, p < 0,05$, celle indiquant l'effet du facteur « Contexte » est $F(4, 5142) = 91, p < 0,05$ et enfin la valeur ANOVA pour l'interaction des deux facteurs « Voyelle*Contexte » est $F(36, 5142) = 36, p < 0,05$. Le résultat est donné dans le Tableau 45.

Source	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle	9	319340165,883	35482240,654	944,445	< 0,0001
contexte	4	13707545,501	3426886,375	91,215	< 0,0001
voyelle*contexte	36	48879167,644	1357754,657	36,140	< 0,0001

Tableau 45 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Contexte et Voyelle*Contexte sur le F3 des voyelles du français (10 locutrices*4 répétitions*3 syllabes)

Nous avons vérifié quels groupes étaient responsables de l'effet combiné du facteur « Voyelle*Contexte » avec le test Fisher. Le résultat complet se trouve dans les annexes à la page 64 et il montre que les moyennes du troisième formant sont significativement différentes de celles de voyelles cibles :

- en contexte labial, pour toutes les voyelles, mis à part le [ø, œ, o, a] (contexte-0 vs contexte-p)
- en contexte dental, pour toutes les voyelles, mis à part le [ɛ, ɔ] (contexte-0 vs contexte-t)
- en contexte palato-vélaire, pour toutes les voyelles, mis à part le [y] (contexte-0 vs contexte-k)
- en contexte uvulaire, pour toutes les voyelles, mis à part le [œ, ɔ, a] (contexte-0 vs contexte-R)

En contexte labial, le F3 des voyelles antérieures étirées [i, e, ɛ] baisse de manière importante (en moyenne de 298 Hz, 324 Hz et 128 Hz respectivement) alors que celui des voyelles antérieures labiales [y, ø, œ] rencontre des variations moindres qui se manifestent par une augmentation de 82 Hz, 32 Hz et 27 Hz respectivement. Ce résultat est attendu car les voyelles labiales s'articulent avec les lèvres arrondies tout comme la consonne [p]. Le F3 de la voyelle [a] n'est pas sensible au contexte labial car il dépend essentiellement de la cavité postérieure.

En contexte dental, le F3 des voyelles antérieures étirées, étant principalement associé à la cavité antérieure, baisse car cette dernière devient plus longue. Cette baisse est particulièrement importante dans le cas du F3 de [i] et [e] ; elle est de 366 Hz et 309 Hz respectivement. En revanche, le F3 des voyelles labiales (naturellement bas) augmente : la montée est en moyenne de 210 Hz pour [y], de 137 Hz pour [ø] et de 104 Hz pour [œ]. Le F3 de la voyelle [a] est de nouveau peu affecté par le contexte dental car il dépend principalement de la cavité postérieure.

Le contexte palato-vélaire provoque une baisse de la valeur du F3 de toutes les voyelles antérieures, à l'exception du [y] qui ne rencontre pas de variations importantes (montée de 27 Hz). Le F3 de [y] dépend en effet essentiellement de la cavité postérieure. La cavité antérieure de [a] en

contexte palato-vélaire est raccourcie au point de déterminer également le F3, comme dans le cas des voyelles antérieures. Ce dernier baisse alors de 140 Hz par rapport à la cible.

Le contexte uvulaire, lui, fait baisser le F3 des voyelles antérieures étirées qui dépend essentiellement de la cavité antérieure, allongée par [ʁ]. Il élève ensuite la valeur du F3 des voyelles antérieures labiales (élévation due au raccourcissement de la cavité antérieure), ce que l'on peut observer dans le spectrogramme de la Figure 73. Le F3 de [a] n'est pas affecté ce qui est attendu car il s'agit d'une résonance essentiellement de la cavité postérieure. Le spectrogramme de la Figure 73 illustre les formants des voyelles antérieures du français [i, e, ε, y, ø, œ] prononcées isolément (iso) et en contexte symétrique uvulaire (R).

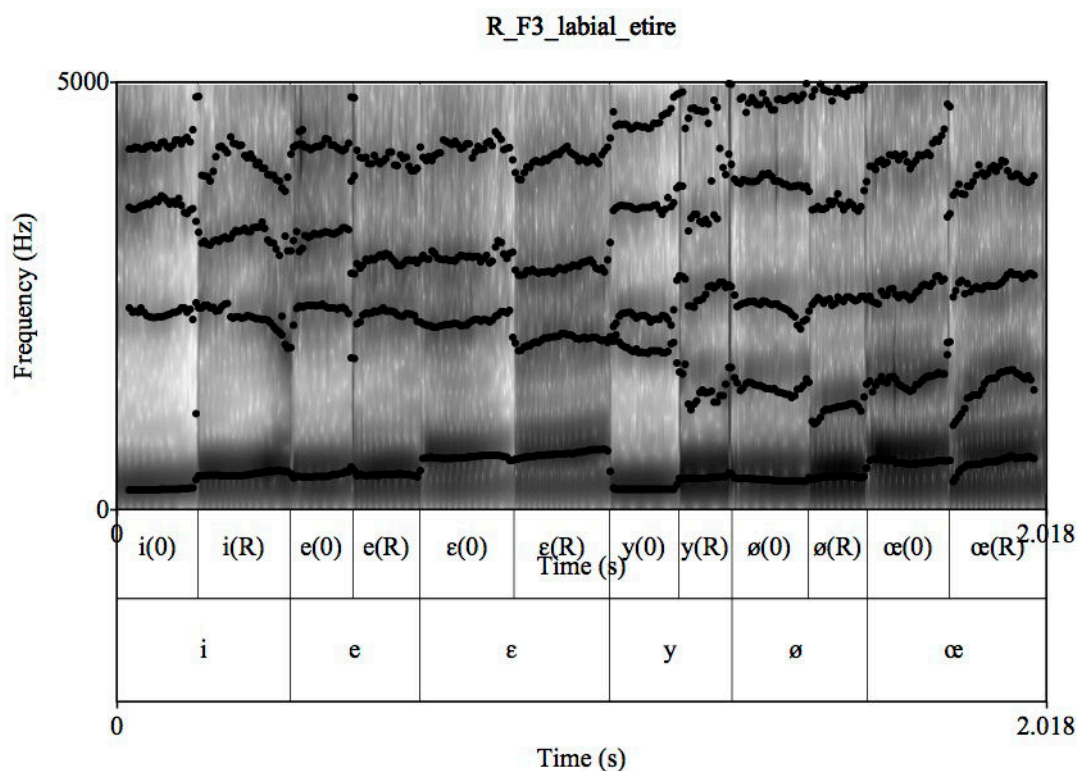


Figure 73 : Spectrogramme des voyelles [i, e, ε, y, ø, œ] du français prononcées par une Française isolément et en contexte RVR dans lequel le F3 des voyelles étirées baisse alors que celui des voyelles labiales augmente

Nous constatons que la consonne [ʁ] provoque une baisse du troisième formant des voyelles étirées, associé à la cavité antérieure qui se prolonge, et qu'elle augmente celui des voyelles labiales dont le [y] qui par conséquent perd son caractère focal. En effet, la diminution d'arrondissement et de protrusion des lèvres lors du contact des voyelles labiales avec une consonne non labiale raccourcit la cavité antérieure, en petite partie responsable du formant F3, ce qui engendre une augmentation de ce formant.

Notre hypothèse de départ, à savoir qu'il existe un effet du contexte phonétique – nul, labial, dental, palato-vélaire, uvulaire - sur les valeurs formantiques F1, F2 et F3 des dix voyelles orales du français [i, e, ε, y, ø, œ, u, o, ɔ, a] est ainsi confirmée et l'hypothèse nulle peut être rejetée.

9.2.4 Coarticulation en français : Quatrième formant

Le quatrième formant est souvent considéré comme la résonance de la cavité laryngale qui forme un tube au-dessus des plis vocaux (Honda *et al.*, 2008). Lorsque le tube se ferme (par fermeture de la glotte), une résonance de type Helmholtz est générée vers 3500 Hz et correspond alors au F4. Ce formant est important pour l'identification de la voyelle [i] qui se réalise prépalatale avec une convergence du troisième et quatrième formant (Vaissière, 2009). Dans la simulation de [i] avec et

sans cavité laryngale, Vaissière a montré dans ses cours de Master que la présence de cette dernière renforçait la région de F4.

Les variations du F4 de [i], dues aux contextes labial, dental et palato-vélaire, (articulations antérieures) sont alors minimales. En revanche, les variations engendrées par le contexte uvulaire sont plus importantes car la cavité antérieure s'allonge considérablement, elle devient alors en partie responsable du formant F4 qui baisse en moyenne de 196 Hz par rapport à la cible.

Les variations du quatrième formant du [i] prononcé par une Française native en isolation et en contextes symétriques p, t, k, R sont illustrées dans le spectrogramme de la Figure 74.

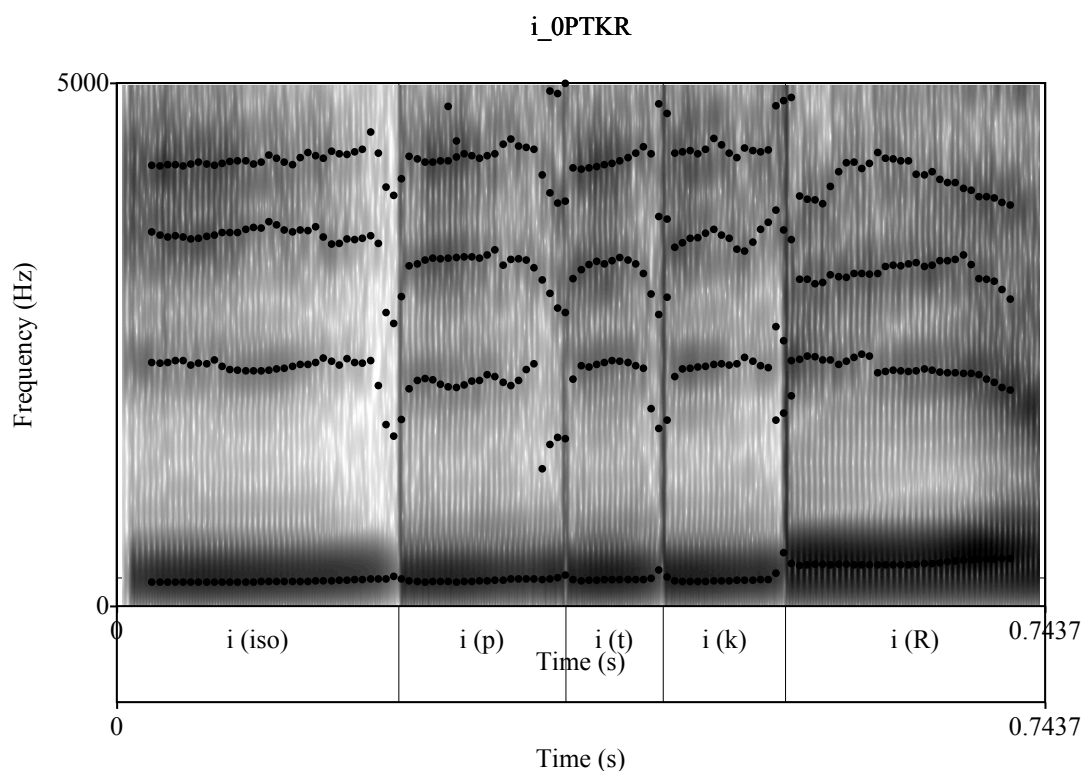


Figure 74 : Spectrogramme de la voyelle [i] du français prononcée par une Française native isolément et en contextes pVp, tVt, kVk, RVR. Notons que seul le contexte uvulaire provoque un mouvement du F4 (qui baisse) par rapport à la cible

9.2.5 Coarticulation en français : Caractère focal

Nous avons étudié le caractère focal des voyelles prononcées en contexte consonantique en calculant la distance acoustique entre les formants (en considérant la valeur centrale) qui sont rapprochés dans le cas des voyelles isolées.

Les distances acoustiques F2-F1 des voyelles postérieures, de F2-F1 et F3-F2 des voyelles antérieures et F2-F1, F3-F2, F4-F3 (en Hertz) de [i], prononcées en isolation et en contextes consonantiques [p, t, k, ʁ] sont indiquées dans le Tableau 46. Ce tableau nous fait remarquer que les formants F2-F1 (voyelles postérieures et [a]), F3-F2 (de [y]) et F4-F3 (de [i]) sont plus rapprochés en isolation qu'en contexte.

Distances	F2-F1					F3-F2					F4-F3				
Voyelle/ contexte	0	p	t	k	R	0	p	t	k	R	0	p	t	k	R
i	2263	2101	2119	2260	2097	1273	1127	1022	1148	795	631	863	905	683	943
y	1768	1747	1786	1769	1314	417	507	587	426	1039					
e	2105	1880	1851	2016	1785	856	734	774	710	735					
ø	1167	1182	1310	1309	858	1045	1045	1031	771	1470					
ɛ	1620	1452	1513	1772	1304	808	832	881	621	957					
œ	998	1027	1190	1275	779	1160	1140	1081	717	1377					
a	470	754	1041	1250	588	1535	1220	1106	713	1370					
u	495	551	918	551	383	1730	1799	1545	1838	2267					
o	362	448	750	495	360	1995	1892	1692	1737	2167					
ɔ	475	588	835	651	448	1818	1532	1421	1347	1780					

Tableau 46 : Distances acoustiques F2-F1 (de toutes les voyelles), F3-F2 (voyelles antérieures) et F4-F3 (voyelle [i]) en Hertz des voyelles en isolation (0) et en contextes p, t, k, R. Les moyennes formantiques calculées à partir de productions de 10 locutrices*4 répétitions (valeur prise à la moitié de la durée vocalique). Le contexte favorisant la focalisation est en gras, le contexte fortement défavorisant est grisé

Comme illustré dans le Tableau 46, certains contextes peuvent « déformer » la voyelle subséquente au point qu'elle perde son caractère focal. Dans les suites [pip], [tit], [kik], la distance acoustique F4-F3 moyenne dépasse 850 Hz et le troisième formant se trouve à mi-chemin entre le deuxième et le quatrième formant. Seul le contexte palato-vélaire conserve le caractère focal proche de la cible, regroupant F4-F3. Le spectrogramme de la Figure 74 de la page 173 montre cet écartement des formants F4-F3 du [i] en contextes [p, t, k] et un rapprochement plus étroit de ces formants en contexte [k] chez une Française native.

Ensuite en contexte uvulaire, la distance acoustique moyenne de F3-F2 du [y] est plus que doublée par rapport à la cible et passe de 417 Hz à 1039 Hz. Cet éloignement des formants est illustrée à la Figure 75 qui est un spectrogramme du [y] français prononcé en isolation et en contextes [k, k] par une Française native. Notons que le contexte palato-vélaire ne modifie pas le caractère focal du [y].

Le caractère focal de la voyelle [a] dans les suites [tat] et [kak] se perd également : la distance F2-F1 passe de 470 Hz en moyenne pour [a] en isolation à 1041 Hz dans [tat] et 1250 Hz dans [kak]. Un fort écartement formantique F2-F1, visible sur le spectrogramme de la Figure 67 de la page 166 concerne également les voyelles [u, o, ɔ] en contexte dental ; la distance F2-F1 est de 918 Hz, 750 Hz et 835 Hz respectivement contre 495 Hz, 362 Hz et 475 Hz pour ces mêmes voyelles en isolation.

En revanche d'autres contextes favorisent le caractère focal, comme nous pouvons le remarquer dans les suites [kik] (avec une distance moyenne F4-F3 de 683 Hz), [kyk] (avec la distance F3-F2 de 426 Hz) et [kuk], [kov], [kɔk] (avec des distances moyennes F2-F1 de 383 Hz, 360 Hz et 448 Hz respectivement). La consonne uvulaire renforce même le caractère focal de ces voyelles postérieures, ce qui est illustré à la Figure 71 de la page 169 qui est un spectrogramme des voyelles postérieures du français, prononcées par une Française native en isolation et en contexte uvulaire.

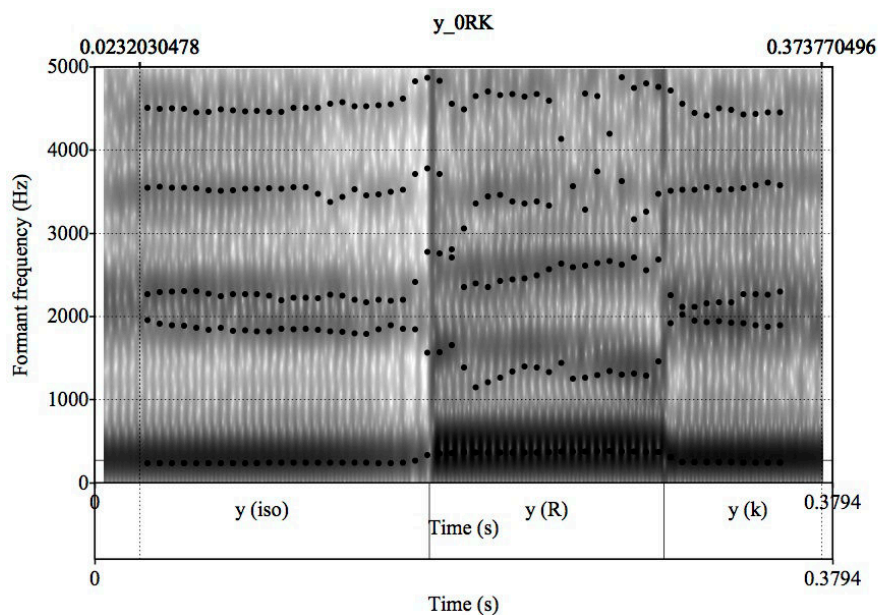


Figure 75 : Spectrogramme du [y] français prononcé par une Française native en contextes nul et [ɣ, k]. Notons que le caractère focal du [y], qui rapproche les formants F3-F2 en contextes nul et [k], est effacé en contexte [ɣ]

Les résultats à propos de l'atténuation du caractère focal des voyelles du français produites en contexte sont en accord avec la littérature. Une comparaison avec les valeurs de Gendrot and Adda-Decker (2005) se trouve dans les annexes (page 65).

9.2.6 Coarticulation en français : Durée vocalique

Puisque la durée vocalique peut affecter les patrons formantiques (Flemming, 2006; Lindblom, 1963), nous avons comparé la durée moyenne des voyelles produites en isolation (4 répétitions*10 locutrices) et en contextes labial, dental, palato-vélaire et uvulaire (4 répétitions*3 positions*10 locutrices). Le résultat se trouve dans le Tableau 47.

Contexte	Isolation	Labial	Dental	Palato-vélaire	uvulaire
Durée moyenne en ms (écart type)	157 (61)	93 (34)	95 (34)	97 (34)	150 (62)

Tableau 47 : Durée moyenne (en ms) des voyelles du français produites en isolation et en contextes p, t, k, R (4 répétitions*3 positions*10 locutrices). L'écart type est entre parenthèses

Le Tableau 47 montre que les voyelles les plus longues sont celles qui ont été produites en isolation (durée moyenne de 157 ms). Il s'agit en même temps de voyelles qui servent de référence dans ce travail et par rapport auxquelles est déterminée l'ampleur de réduction des voyelles produites en contextes p, t, k, R. Ainsi, la réduction vocalique observée dans ce chapitre n'est pas seulement due au contexte phonétique qui « perturbe » les valeurs prototypiques des voyelles produites en isolation mais également à la durée qui est d'environ 30 % moins importante dans le cas des voyelles produites en contexte labial (la durée moyenne est de 93 ms en moyenne), dental (de 95 ms) et palato-vélaire (de 97 ms). En revanche, puisque le contexte uvulaire a un effet allongeant (la durée vocalique moyenne est de 150 ms), la réduction phonétique des voyelles insérées dans ce contexte s'explique uniquement par l'influence du contexte et non pas par une durée plus petite.

9.2.7 Coarticulation en français : Conclusion partielle

Le premier formant est moins variable que le deuxième formant. Les variations du deuxième formant sont systématiques et plus ou moins importantes selon le contexte. *La plage de variation est plus importante avec les voyelles postérieures.* Les variations du troisième formant sont les plus importantes pour [i], puis pour [e] où F3 est essentiellement affilié à la cavité antérieure.

Le Tableau 48 récapitule les mouvements des deuxième et troisième formants dus aux contextes p, t, k, R calculés à partir de nos données réelles, pour quatre groupes de voyelles : antérieures étirées/ arrondies, postérieures et la voyelle [a] (postérieure en isolation et centrale/ antérieure en contexte). La flèche vers le haut montre que la valeur du formant par rapport à la cible augmente, la flèche vers le bas montre qu'elle baisse. Le nombre de flèches illustre l'ampleur des mouvements formantiques (une flèche est égale à env. 100 Hz).

Formants	F2				F3			
Voyelle/contexte	p	t	k	R	p	t	k	R
Antérieures étirées	↓	↓	-	↓	↓↓	↓↓	↓	↓↓
Antérieures arrondies	--	↑	↑	↓↓	--	↑	↓ sauf [y]	↑
Postérieures	↑	↑↑↑	↑	--	/	/	/	/
[a]	↑↑	↑↑↑	↑↑↑↑	↑	--	--	↓	--

Tableau 48 : Les mouvements de F2 et F3 des voyelles antérieures étirées/ arrondies, voyelles postérieures et [a] en contextes p, t, k, R par rapport à la cible. ↑ signifie une augmentation de la valeur et ↓ indique sa baisse. La case grisée indique un résultat inattendu selon la modélisation dans VTDemo. Le nombre de flèches indique l'ampleur des mouvements formantiques (1 flèche par tranche de 100 Hz)

Le Tableau 49 montre ensuite l'ampleur des mouvements formantiques selon le degré de constriction de chaque voyelle antérieure.

Formants	F2				F3			
Voyelle/contexte	p	t	k	R	p	t	k	R
i	↓	↓	--	--	↓↓↓	↓↓↓	↓	↓↓↓↓↓
e	↓↓	↓↓	--	↓↓	↓↓↓	↓↓↓	↓↓	↓↓↓
ɛ	↓	↓	--	↓↓	↓	--	(↓)	(↓)
y	--	--	--	↓↓↓	--	↑↑	--	↑↑
ø	--	↑	↑	↓↓	--	↑+	↓	↑
œ	--	↑+	↑↑	↓	--	↑	↓↓	--

Tableau 49 : Les mouvements de F2 et F3 des voyelles antérieures étirées et arrondies en contextes [p, t, k, ʁ] par rapport à la cible. ↑ signifie une augmentation de la valeur et ↓ indique sa baisse. Le nombre de flèches indique l'ampleur des mouvements formantiques (1 flèche par tranche de 100 Hz)

Les Tableau 48 et Tableau 49 mettent en évidence que le contexte labial provoque par rapport à la cible :

- une baisse du F2 et F3 des voyelles antérieures étirées
- une élévation du F2 des voyelles postérieures proportionnellement au degré d'aperture (augmentation maximale pour les voyelles mi-ouvertes)
- une élévation de F2 de [a]

Le contexte dental engendre ensuite :

- une baisse du F2 et F3 des voyelles antérieures étirées
- une élévation du F2 des voyelles labiales (et de F3 des voyelles antérieures)
- une forte élévation de F2 du [a]

Le contexte palato-vélaire provoque les mouvements formantiques suivants :

- une baisse du F3 des voyelles antérieures étirées
- une élévation du F2 des voyelles postérieures et une forte élévation de F2 du [a]
- une élévation du F2 et une baisse du F3 des voyelles antérieures labiales
- pas de mouvement pour [y] qui résiste à la coarticulation

Le contexte uvulaire affecte les formants par rapport à la cible de la façon suivante :

- une baisse du F2 et F3 des voyelles antérieures étirées
- une baisse du F2 et une élévation du F3 des voyelles antérieures labiales
- une élévation du F2 du [a]

En reprenant le terme de compatibilité de van Bergem (1993), les sons du français *compatibles* sont les suivants :

Consonne	Compatibilité	Voyelle(s)
R	postérieures	u, o, ɔ
p	arrondies	y, ø, œ, u, o
k	antérieures fermées :	i, y

Tableau 50 : Contextes consonantiques qui conservent les valeurs formantiques cibles de voyelles du français subséquentes

Notons également que la réduction vocalique observée en français n'est pas seulement occasionnée par l'effet du contexte mais également par la durée de la voyelle, ce qui est en accord avec l'étude de Lindblom (1963). En effet, la durée vocalique est sujette à de grandes variations en français. En moyenne, elle est de 106 ms pour les voyelles issues de logatomes et de 157 ms pour les voyelles isolées qui servent de référence.

9.3 Coarticulation en tchèque : Données réelles

Afin d'étudier les variations acoustiques des voyelles monophthongues du tchèque prononcées en contextes consonantiques [p, t, k, h] par rapport à la cible, nous avons d'abord calculé les moyennes formantiques de chaque voyelle (valeurs prises à la moitié de la durée vocalique, toute position du mot confondue). Les moyennes avec les écarts types se trouvent dans les annexes (à la page 66). La construction du corpus et le calcul des moyennes sont récapitulés dans le Tableau 51.

CORPUS
10 voyelles monophthongues du tchèque [ɪ, i:, ɛ, ɛ:, u, u:, o, o:, a, a:], insérées dans des logatomes pVpVpVp, tVtVtVt, kVkVkVk, hVhVhVh d'une phrase cadre, exemple : « Slovo papapap pŭsobí divně ¹⁷ . », répétées 4 fois
LOCUTRICES
20 Tchèques natives de la région de Bohême
CALCUL DE LA MOYENNE
A partir de la valeur relevée à la moitié de la durée de chaque voyelle monophthongue des logatomes trissyllabiques (20 locutrices* 4 répétitions* 3 syllabes)

Tableau 51 : Corpus, locutrices et mesures des moyennes formantiques des voyelles du tchèque prononcées en contextes p, t, k, h

Le Tableau 52 indique les changements de valeurs formantiques moyennes (en Hz) des voyelles en contextes [p, t, k, h] par rapport aux valeurs formantiques des voyelles isolées (la cible). Les chiffres positifs signifient une élévation des formants, les chiffres négatifs indiquent leur baisse.

<p>Dans l'étude des effets de coarticulation, nous émettons l'hypothèse que les valeurs formantiques des voyelles monophthongues du tchèque prononcées en contextes [p, t, k, h] sont différentes de celles des voyelles prononcées en isolation.</p> <p><i>L'hypothèse nulle consiste à dire qu'il n'existe pas de différences significatives entre les valeurs formantiques des voyelles prononcées en isolation et en contexte consonantique.</i></p>
--

Pour mesurer l'effet du contexte sur les formants F1, F2 et F3 des dix voyelles du tchèque, nous avons conduit une analyse ANOVA à deux facteurs : voyelle et contexte. Le facteur « Voyelle » possède dix modalités qui sont les voyelles [ɪ, i:, ɛ, ɛ:, o, o:, u, u:, a, a:] et le facteur « Contexte » possède cinq modalités, à savoir les contextes nul, labial, dental, palato-vélaire et glottal. Nous nous sommes essentiellement intéressée à l'effet combiné des deux facteurs « Voyelle*Contexte » afin de vérifier comment évoluent les valeurs formantiques de chacune des dix voyelles dans les cinq contextes phonétiques. Les résultats de l'analyse de variance sont donnés pour chaque formant séparément.

¹⁷ Nous traduisons « Le mot papapap donne une impression bizarre. »

	F1				F2				F3				F4			
	p	t	k	h	p	t	k	h	p	t	k	h	p	t	k	h
i	16	2	-42	-13	-279	-281	52	-45	-188	-116	16	-42	-65	42	30	-8
i:	5	11	6	23	-59	-49	3	-30	-73	-104	-33	-130	-3	-21	-18	-34
ɛ	-75	-124	-192	-16	-171	-130	219	-16	-66	15	-95	1				
ɛ:	1	-42	-92	2	-104	-99	117	-38	6	31	-46	6				
a	-104	-115	-114	-33	-34	121	184	20	58	149	-198	80				
a:	-37	-45	-51	-35	39	141	175	35	30	54	-145	58				
u	8	-11	-5	10	43	310	52	99								
u:	6	2	10	13	69	248	79	91								
o	-69	-62	-65	-45	-10	162	48	50								
o:	10	9	8	10	32	118	72	75								

Tableau 52 : Distance acoustique (en Hertz) entre les valeurs moyennes formantiques centrales des voyelles tchèques prononcées en contextes [p, t, k, fi], calculées à partir de productions de 20 locutrices*4 répétitions*3 syllabes et les valeurs formantiques moyennes des mêmes voyelles prononcées en isolation, calculées à partir de productions de 20 locutrices* 4 répétitions*3 valeurs. Les mouvements formantiques importants sont grisés

9.3.1 Coarticulation en tchèque : Premier formant

Comme l'indique le Tableau 52, les variations du premier formant des voyelles fermées [i, i:, u, u:] et de la voyelle moyenne [o:] sont inférieurs à 50 Hz par rapport à la cible. En revanche, le F1 des voyelles brèves, de grande ou moyenne aperture [ɛ, o, a] en contextes [p, t, k] et des voyelles longues [ɛ:, a:], en contexte [k], baisse de plus de 50 Hz. Cet abaissement est particulièrement important quand la voyelle [ɛ] est prononcée en contextes [t, k] (F1 baisse respectivement de 124 Hz et 192 Hz) et quand le [a] se trouve en contextes [p, t, k] (F1 baisse respectivement de 104 Hz, 115 Hz et 114 Hz). L'effet du contexte palato-vélaire provoquant une baisse importante du premier formant des voyelles [ɛ, ɛ:, a] corrobore l'étude de Hillenbrand *et al.* (2001) sur les voyelles de l'anglais. Ce phénomène est illustré sur le spectrogramme de la Figure 76 qui compare ces voyelles prononcées par une Tchèque native en isolation (iso) et en contexte palato-vélaire (k).

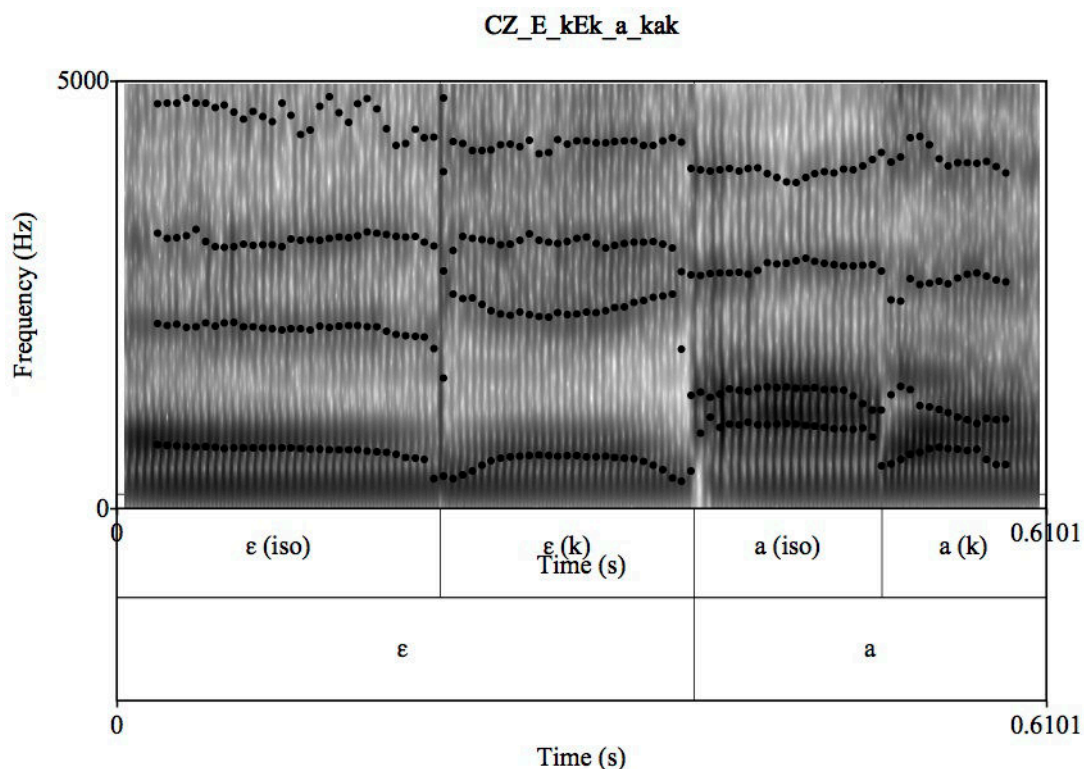


Figure 76 : Spectrogramme des voyelles [ε, a] du tchèque prononcées par une Tchèque isolément et en contexte kVk qui provoque une baisse de F1. Notons que le [a] en isolation est postérieur

Le graphe de la Figure 77 illustre les variations du premier formant (valeur prise à la moitié de la durée vocalique) des dix voyelles monophtongues du tchèque [ɪ, i:, ε, ε:, o, o:, u, u:, a, a:] prononcées isolément (0) et en contextes [p, t, k, h].

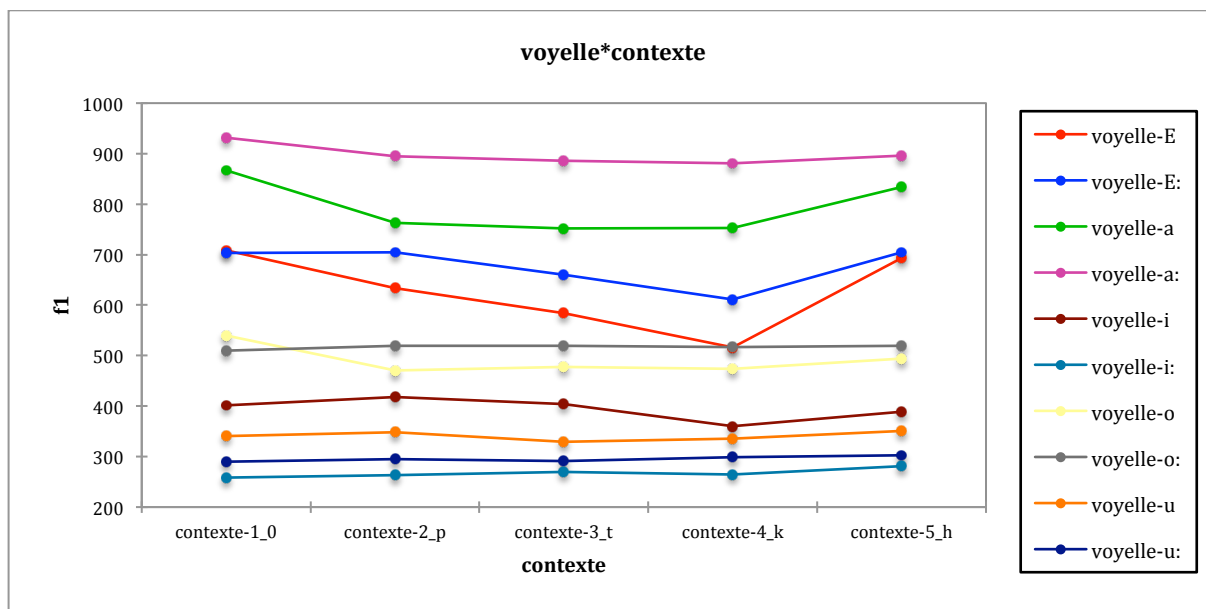


Figure 77 : Le F1 moyen des dix voyelles monophtongues du tchèque [ɪ, i:, ε, ε:, o, o:, u, u:, a, a:] prononcées isolément (0) et en contextes [p, t, k, h] dans des logatomes CVCVCVC (20 locutrices*4 répétitions)

La Figure 77 fait ressortir que le contexte glottal affecte la valeur du premier formant minimalement, comme attendu.

Le résultat de l'analyse de variance exprimant l'effet global du contexte sur la valeur formantique F1 des dix voyelles est indiqué dans le Tableau 53. La valeur ANOVA est $F(49, 10349) = 1986, p < 0,05$.

F1	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	49	427083158,185	8715982,820	1985,593	< 0,0001
Erreur	10349	45428104,809	4389,613		

Tableau 53 : Résultat de l'ANOVA à 2 facteurs qui mesure l'effet du contexte sur le formant F1 des dix voyelles monophthongues du tchèque (20 locutrices*4 répétitions*3 syllabes)

Le Tableau 54 indique la valeur ANOVA du facteur « Voyelle » qui est $F(9, 10349) = 10580, p < 0,05$, ainsi que du facteur « Contexte » où $F(4, 10349) = 192, p < 0,05$ et enfin de l'interaction des deux facteurs « Voyelle*Contexte » avec $F(36, 10349) = 36, p < 0,05$.

F1	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle	9	417973651,323	46441516,814	10579,866	< 0,0001
contexte	4	3378542,491	844635,623	192,417	< 0,0001
voyelle*contexte	36	5730964,372	159193,455	36,266	< 0,0001

Tableau 54 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Contexte et Voyelle*Contexte sur le F1 des voyelles du tchèque (20 locutrices*4 répétitions*3 syllabes)

Afin de vérifier à quel groupe est dû l'effet global combiné des deux facteurs « Voyelle*Contexte » sur le F1, nous avons conduit le test a posteriori de Fisher dont les résultats complets se trouvent dans les annexes (page 66). Ils montrent que les valeurs moyennes formantiques des voyelles en isolation et celles des voyelles en contexte sont significativement différentes au niveau du premier formant :

- en contexte labial, pour les voyelles [ɛ, a, a:, o] (contexte-0 vs contexte-p)
- en contexte dental, pour les voyelles [ɛ, ɛ:, a, a:, o] (contexte-0 vs contexte-t)
- en contexte palato-vélaire, pour les voyelles [ɪ, i:, ɛ, ɛ:, a, a:, o] (contexte-0 vs contexte-k)
- en contexte glottal, pour les voyelles [ɪ, i:, ʌ, a:, o] (contexte-0 vs contexte-h)

9.3.2 Coarticulation en tchèque : Deuxième formant

En tchèque, les variations du deuxième formant (valeur prise à la moitié de la durée vocalique) des dix voyelles monophthongues du tchèque [ɪ, i:, ɛ, ɛ:, o, o:, u, u:, a, a:] prononcées isolément (0) et en contextes [p, t, k, h] sont plus importantes que celles du premier formant, comme en français. Le graphe de la Figure 78 illustre ces variations.

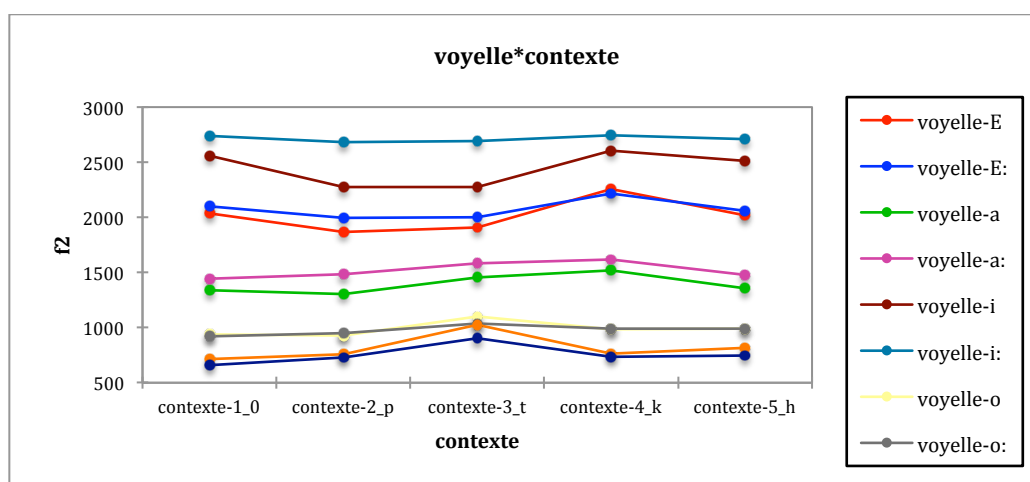


Figure 78 : Le F2 moyen des dix voyelles monophthongues du tchèque [i, i:, ε, ε:, o, o:, u, u:, a, a:] prononcées isolément (0) et en contextes [p, t, k, h] dans des logatomes CVCVCVC (20 locutrices*4 répétitions)

Nous remarquons à la Figure 78 que le contexte palato-vélaire engendre une augmentation du F2 de toutes les voyelles et le contexte dental élève le F2 des voyelles postérieures. La voyelle [i:] ne subit pas de variations de valeurs formantiques importantes.

Le résultat du test ANOVA indique que l'effet global du contexte phonétique sur le deuxième formant des dix voyelles du tchèque est significatif. La valeur ANOVA, renseignée dans le Tableau 55, est $F(49, 10349) = 4827$ avec $p < 0,05$.

F2	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	49	4628631910,539	94461875,725	4826,544	< 0,0001
Erreur	10349	202543658,574	19571,327		

Tableau 55 : Résultat de l'ANOVA à 2 facteurs qui mesure l'effet du contexte sur le formant F2 des dix voyelles monophthongues du tchèque (20 locutrices*4 répétitions*3 syllabes)

La variable nominale « Voyelle » a un effet sur le formant F2 qui est exprimé par la valeur ANOVA $F(9, 10349) = 25779$, $p < 0,05$. La valeur ANOVA de l'effet global du facteur « Contexte » est $F(4, 10349) = 363$, $p < 0,05$ et celle de l'effet combiné des facteurs « Voyelle*Contexte » est $F(36, 10349) = 84$, $p < 0,05$ (Tableau 56).

F2	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle	9	4540759932,566	504528881,396	25778,982	< 0,0001
contexte	4	28422942,940	7105735,735	363,069	< 0,0001
voyelle*contexte	36	59449035,033	1651362,084	84,377	< 0,0001

Tableau 56 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Contexte et Voyelle*Contexte sur le F2 des voyelles du tchèque (20 locutrices*4 répétitions*3 syllabes)

Nous avons examiné à quels groupes est dû l'effet global combiné des facteurs « Voyelle*Contexte » par le test a posteriori de Fisher, dont les résultats complets sont dans les annexes (page 67). Ils montrent que les valeurs moyennes du deuxième formant sont significativement différentes de celles des voyelles isolées :

- en contexte labial, pour toutes les voyelles, mis à part le [a, o, o:] (contexte-0 vs contexte-p)

- en contexte dental, pour toutes les dix voyelles (contexte-0 vs contexte-t)
- en contexte palato-vélaire, pour toutes les voyelles, mis à part le [i:] (contexte-0 vs contexte-k)
- en contexte glottal, pour [ɪ, ɛ:, u, u:, o, o:] (contexte-0 vs contexte-h)

9.3.2.1 Effet du contexte labial sur le F2 (en tchègue)

Le triangle F1/F2 (en Bark) montre l'évolution de la valeur du deuxième formant des voyelles du tchègue entre le contexte isolée servant de référence (trait plein) et le contexte symétrique labial (traits pointillés). L'écart type tracé est de un.

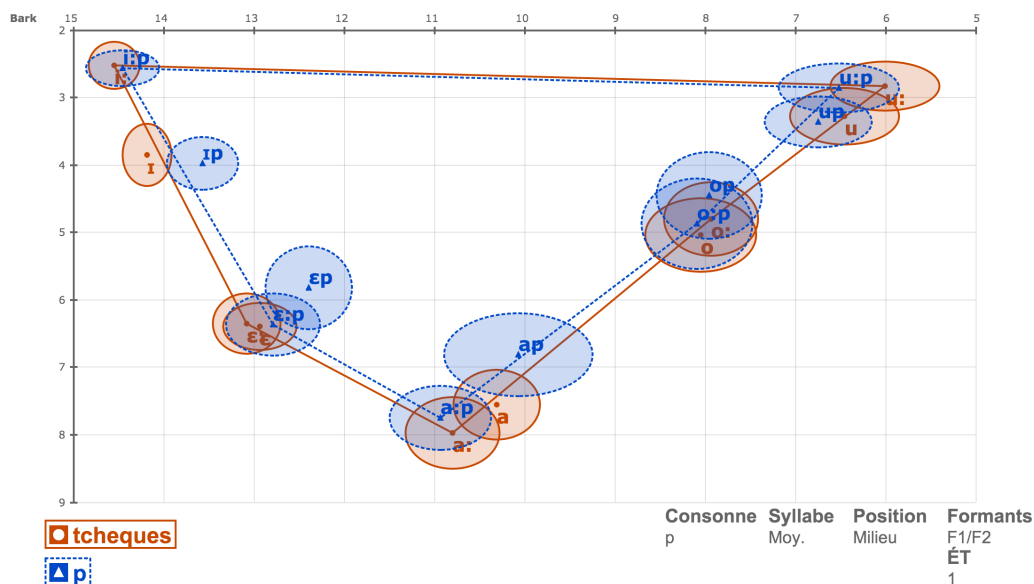


Figure 79 : Triangle vocalique des voyelles monophthongues du tchègues en isolation (trait plein) et en contexte symétrique labial (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 20 Tchègues *4 répétitions de chaque voyelle* valeur prise à la moitié de la durée vocalique

Les résultats du Tableau 52 de la page 179 et de la Figure 79 montrent que le contexte labial provoque, par rapport à la valeur cible, une baisse de plus de 100 Hz de la valeur du deuxième formant des voyelles antérieures [ɪ, ɛ, ɛ:] qui est essentiellement affilié à la cavité antérieure. Cet abaissement de F2 est particulièrement important dans le cas de la voyelle fermée [ɪ] (de 279 Hz en moyenne) et de [ɛ] (de 130 Hz) qui sont brèves. Plus la voyelle s'allonge, moins la valeur du deuxième formant est affectée par le contexte labial, comme attendu. Les voyelles postérieures sont marquées par une légère augmentation de F2 au contact de la consonne [p], mis à part la voyelle [o] dont la valeur baisse de 10 Hz (mouvement négligeant).

9.3.2.2 Effet du contexte dental sur le F2 (en tchègue)

Le triangle F1/F2 (en Bark) montre l'évolution de la valeur du deuxième formant des voyelles du tchègue entre le contexte isolée servant de référence (trait plein) et le contexte symétrique dental (traits pointillés). L'écart type tracé est de un.

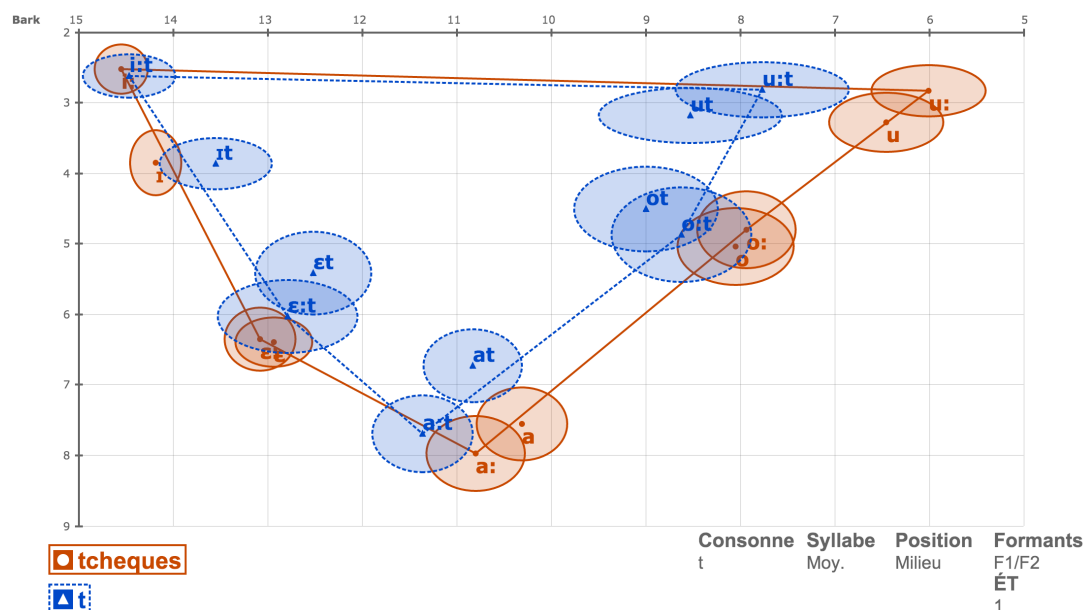


Figure 80 : Triangle vocalique des voyelles monophtongues du tchèque en isolation (trait plein) et en contexte symétrique dental (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 20 Tchèques *4 répétitions de chaque voyelle*valeur prise à la moitié de la durée vocalique

Le contexte dental (Figure 80) engendre une baisse de la valeur du deuxième formant des voyelles antérieures, et en particulier celle de la voyelle fermée brève [ɪ] (de 281 Hz) et une élévation du deuxième formant des voyelles postérieures, notamment celui des voyelles [u], [u:] (qui s'élève respectivement de 310 Hz et 248 Hz en moyenne). Le F2 de la voyelle [a], qui dépend de la cavité antérieure et postérieure, augmente de 121 Hz en moyenne et celui de [a:] long de 141 Hz en moyenne.

9.3.2.3 Effet du contexte palato-vélaire sur le F2 (en tchèque)

Le triangle vocalique de la Figure 81 illustre la valeur du deuxième formant des voyelles du tchèque dans le contexte isolé servant de référence (trait plein) et le contexte symétrique palato-vélaire (traits pointillés). L'écart type tracé est de un.

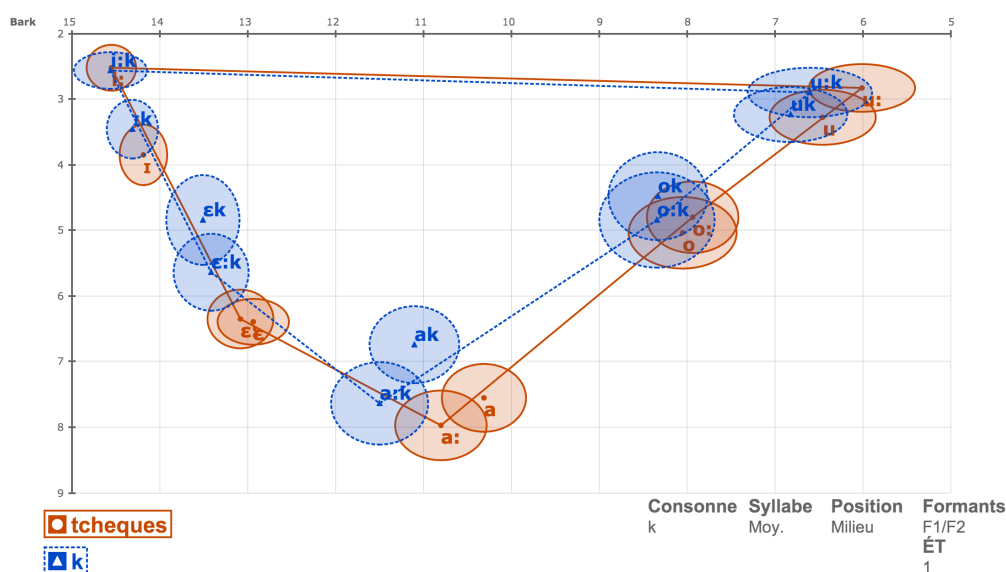


Figure 81 : Triangle vocalique des voyelles monophtongues du tchèque en isolation (trait plein) et en contexte symétrique palato-vélaire (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 20 Tchèques *4 répétitions de chaque voyelle*valeur centrale

Le contexte palato-vélaire provoque une augmentation du deuxième formant de toutes les voyelles qu'il englobe, comme le montre le triangle vocalique F1/F2 de la Figure 81. Il affecte maximale-ment le F2 des voyelles [ɛ, ɛ:] qui s'élève de 219 Hz et 117 Hz respectivement mais également celui des voyelles [a, a:] qui s'élève de 184 Hz et 175 Hz respectivement. Le F2 des autres voyelles augmentent de moins de 80 Hz.

9.3.2.4 Effet du contexte glottal sur le F2 (en tchèque)

Enfin, le triangle vocalique de la Figure 82 illustre le deuxième formant des voyelles du tchèque dans le contexte isolé servant de référence (trait plein) et en contexte symétrique glottal (traits pointillés). L'écart type tracé est de un.

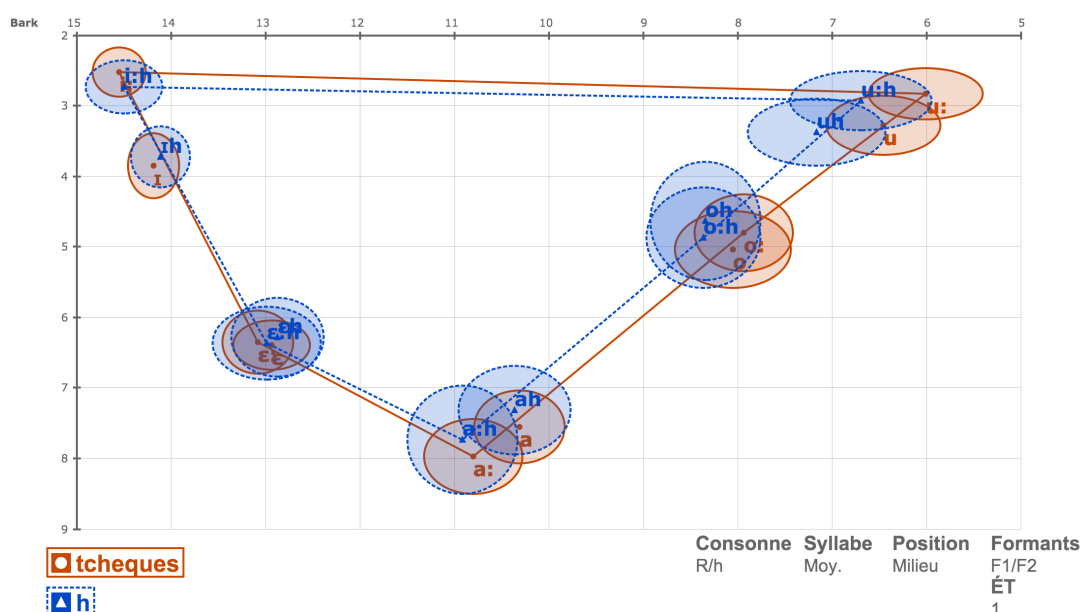


Figure 82 : Triangle vocalique des voyelles monophtongues du tchèque en isolation (trait plein) et en contexte symétrique glottal (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 20 Tchèques * 4 répétitions de chaque voyelle * valeur prise à la moitié de la durée vocalique

Le contexte glottal (Figure 82) provoque une légère chute du deuxième formant des voyelles antérieures, qui baisse de moins 50 Hz en moyenne (mouvement négligeable), et une légère montée du F2 des voyelles postérieures, notamment celui des voyelles [u, u:, o:] (respectivement de 99 Hz, 91 Hz et 75 Hz).

Le deuxième formant de la voyelle [i:] étant affilié à la cavité postérieure ne subit pas de variations importantes dues au contexte phonétique glottal, ni à d'autres contextes.

9.3.3 Coarticulation en tchèque : Troisième formant

L'influence du contexte sur le troisième formant a été étudiée uniquement pour les voyelles antérieures et pour le [a], comme expliqué dans la partie méthodologique. Le Tableau 52 de la page 179 montre que le contexte labial affecte d'une manière importante le troisième formant de la voyelle antérieure fermée brève [ɪ] (qui dépend de la cavité antérieure et postérieure) dont la valeur baisse de 188 Hz par rapport à la cible. Le troisième formant des autres voyelles change de moins de 80 Hz en moyenne, ce qui est négligeable dans le cas du F3.

Le contexte dental provoque une baisse du troisième formant des voyelles [ɪ, i:] de 116 Hz et 104 Hz en moyenne et une montée de 149 Hz de la voyelle [a]. Le F3 dépend essentiellement de la

cavité antérieure dans le cas de [i:] long et en partie de la cavité antérieure dans le cas de [ɪ] bref. Les mouvements formantiques de [ɛ, ɛ:, a:] dont le F3 dépend essentiellement de la cavité postérieure sont minimes.

Le contexte palato-vélaire provoque en général une baisse du troisième formant à l'exception du [ɪ], dont la valeur augmente de 16 Hz (mouvement formantique minime). Le changement le plus important concerne la voyelle ouverte [a] dont la valeur baisse de 198 Hz (le F3 dépend de la cavité antérieure et postérieure dans le cas de [a] bref dans [kak]) et un abaissement de 145 Hz est également observé dans le cas de [a:] long.

Enfin, le contexte glottal ne modifie pas le troisième formant d'une manière importante (les changements formantiques sont inférieurs à 100 Hz), à l'exception du [i:] dont le F3 baisse de 130 Hz, ce qui est contre les attentes.

Le graphe de la Figure 83 montre l'évolution de la valeur du troisième formant due au contexte pour chaque voyelle séparément.

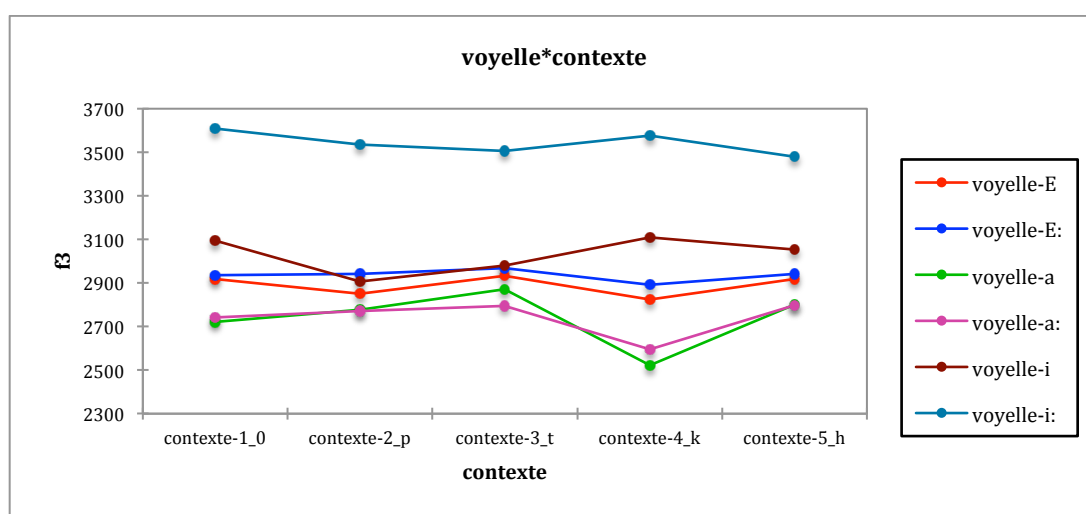


Figure 83 : Le F3 moyen des voyelles monophtongues du tchèque [ɪ, i:, ɛ, ɛ:, a, a:] prononcées isolément (0) et en contextes [p, t, k, ɦ] dans des logatomes CVCVCVC (20 locutrices*4 répétitions)

La Figure 83 représente le F3 des voyelles monophtongues du tchèque [ɪ, i:, ɛ, ɛ:, a, a:] prononcées isolément (0) et en contextes [p, t, k, ɦ] : l'on observe une baisse particulière du troisième formant des voyelles [a, a:] et une hausse du F3 du [ɪ] en contexte palato-vélaire. Les autres contextes affectent peu le troisième formant des six voyelles étudiées.

Le résultat de l'analyse de variance concernant l'effet global du contexte phonétique sur le troisième formant des dix voyelles du tchèque est donné dans le Tableau 57 par la valeur ANOVA $F(49, 10349) = 273, p < 0,05$.

F3	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	49	599217857,738	12228935,872	272,912	< 0,0001
Erreur	10349	463728750,699	44809,040		

Tableau 57 : Résultat de l'ANOVA à 2 facteurs qui mesure l'effet du contexte sur le formant F3 des dix voyelles monophtongues du tchèque (20 locutrices*4 répétitions*3 syllabes)

L'effet du facteur « Voyelle » sur le F3 est significatif car la valeur ANOVA est $F(9, 10349) = 1327, p < 0,05$. La valeur ANOVA exprimant l'effet du facteur « Contexte » est $F(4, 10349) = 129, p < 0,05$ et celle qui renseigne sur l'effet combiné des deux facteurs « Voyelle*Contexte » est $F(36, 10349) = 25, p < 0,05$ (Tableau 58).

F3	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle	9	535264456,136	59473828,460	1327,273	< 0,0001
contexte	4	23204592,095	5801148,024	129,464	< 0,0001
voyelle*contexte	36	40748809,507	1131911,375	25,261	< 0,0001

Tableau 58 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Contexte et Voyelle*Contexte sur le F3 des voyelles du tchèque

Nous avons ainsi obtenu un effet global de l'interaction des facteurs « Voyelle*Contexte » mais il n'est pas dû à tous les groupes. Pour savoir à quels groupes cet effet général est dû, nous avons effectué le test a posteriori de Fisher dont les résultats se trouvent dans les annexes (page 68). Ils montrent que le F3 moyen des voyelles en contexte est significativement différent de celui des voyelles isolées :

- en contexte labial, pour toutes les voyelles, mis à part le [ɛ:, a:, o:] (contexte-0 vs contexte-p)
- en contexte dental, pour toutes les voyelles mis à part le [ɛ, ɛ:] (contexte-0 vs contexte-t)
- en contexte palato-vélaire, pour toutes les voyelles, mis à part le [ɪ, i:, ɛ:, u] (contexte-0 vs contexte-k)
- en contexte glottal, pour toutes les voyelles mis à part le [ɪ, ɛ, ɛ:, u] (contexte-0 vs contexte-h)

Pour résumer, l'hypothèse que nous avons émise au départ est alors confirmée : il existe un effet significatif du contexte phonétique, qui est nul, labial, dental, palato-vélaire ou glottal sur les valeurs formantiques F1, F2, F3 des dix voyelles monophthongues du tchèque [ɪ, i:, ɛ, ɛ:, o, o:, u, u:, a, a:] et l'hypothèse nulle est rejetée.

9.3.4 Coarticulation en tchèque : Quatrième formant

Le quatrième formant des voyelles [ɪ, i:] ne subit pas de variations importantes au contact des consonnes [p, t, k, ɦ] (les mouvements formantiques sont inférieurs à 70 Hz), (voir Tableau 52 de la page 179). Nous pouvons observer ce phénomène sur le spectrogramme de la voyelle [ɪ] (Figure 84), prononcée par une Tchèque native en isolation et en contextes labial, dental, palato-vélaire et glottal. Il est à noter que la valeur du F4 change très peu selon le contexte dans lequel la voyelle avait été prononcée.

Notons que l'absence de variations du F4 de [ɪ] tchèque prononcé dans les contextes phonétiques symétriques p, t, k, h par rapport à la cible coïncide avec les résultats sur les mouvements du F4 du [i] français (voir Figure 74) où seule l'insertion en contexte uvulaire provoque un abaissement du quatrième formant.

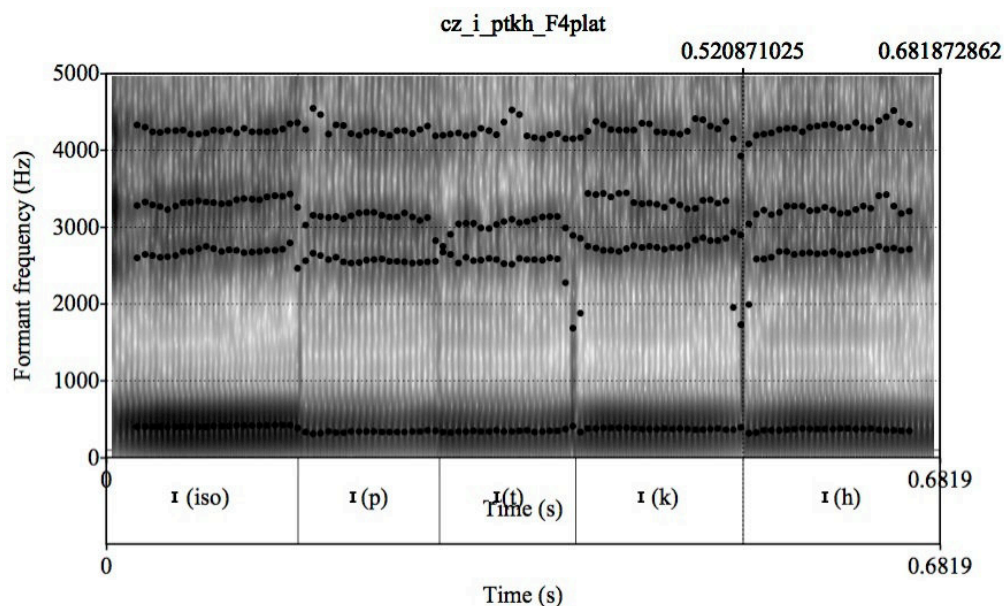


Figure 84 : Spectrogramme de la voyelle [ɪ] du tchèque prononcée par une Tchèque isolément et en contextes pVp, tVt, kVk, hVh. Notons qu'aucun contexte n'engendre de modifications importantes du F4

9.3.5 Coarticulation en tchèque : Caractère focal

De même qu'en français, les distances acoustiques entre les formants rapprochés des voyelles focales sont réduites quand ces dernières se trouvent en isolation.

	F2-F1					F3-F2					F4-F3				
	0	p	t	k	h	0	p	t	k	h	0	p	t	k	h
ɪ	2154	1859	1872	2248	2122	539	630	703	503	542	1091	1213	1249	1104	1125
i:	2483	2419	2423	2480	2430	868	854	812	832	768	651	722	735	667	747
ɛ	1328	1232	1322	1739	1328	881	986	1025	568	898					
ɛ:	1395	1290	1339	1604	1356	839	949	968	676	882					
a	469	540	705	768	523	1385	1476	1413	1002	1445					
a:	511	587	698	738	581	1298	1289	1211	977	1320					
u	374	409	696	431	463	2023	2047	1846	1976	1966					
u:	369	432	614	438	447	1970	2027	1877	1985	2016					
o	399	457	623	513	495	1816	1865	1874	1624	1845					
o:	408	431	517	473	473	1836	1896	1890	1679	1876					

Tableau 59 : Distances acoustiques F2-F1 (de toutes les voyelles), F3-F2 (voyelles antérieures) et F4-F3 (voyelles [ɪ, i:]) en Hertz des voyelles du tchèque selon le contexte phonétique. Les moyennes formantiques sont calculées à partir de productions de 20 locutrices*4 répétitions (valeur prise à la moitié de la durée vocalique). Le contexte favorisant la focalisation est en gras, le contexte fortement défavorisant est grisé

Comme l'indique le Tableau 59, les formants F2-F1 des voyelles postérieures et du [a] et F3-F4 du [i:] s'éloignent systématiquement quand les voyelles sont prononcées en contexte. Le contexte dental [t] provoque une séparation importante des formants F2-F1 des voyelles postérieures et la consonne palato-vélaire [k] écarte les formants F2-F1 des voyelles [a] et [a:]. Ainsi, les formants F2-F1 des voyelles des suites [tat], [kak], [ta:t], [ka:k], [tut], [tu:t] et [tot] sont écartés d'environ 200 Hz de plus par rapport à ceux des voyelles isolées. La voyelle isolée [a] est caractérisée par un rapprochement F2-F1 de moins de 500 Hz. Dès qu'elle est prononcée en contextes [t, k], son F2 augmente et par conséquent un écartement plus important des deux premiers formants se produit. La voyelle longue [a:] est caractérisée par une distance acoustique F2-F1 supérieure à 500 Hz en isolation et dans tous les contextes. La variation de la distance F4-F3 de la voyelle [i:] selon le contexte est inférieure à 100 Hz : elle est de 651 Hz pour [i:] en isolation, et passe à 722 Hz, 735 Hz, 667 Hz, 747 Hz en contextes [p, t, k, h] respectivement. Comme attendu, la partie stable des voyelles longues est moins sujette aux variations dues au contexte phonétique que celle des voyelles brèves (Lindblom, 1963).

Les résultats à propos de l'atténuation du caractère focal des voyelles du tchèque produites en contexte sont en accord avec la littérature. Une comparaison avec les valeurs de Skarnitzl and Volin (2012) se trouve dans les annexes (page 69).

9.3.6 Coarticulation en tchèque : Durée vocalique

Le contexte consonantique peut avoir également un effet sur la durée vocalique. Le Tableau 60 compare la durée moyenne des voyelles produites en isolation (4 répétitions*20 locutrices) et en contextes labial, dental, palato-vélaire et glottal (4 répétitions*3 positions*20 locutrices).

Contexte	Isolation	Labial	Dental	Palato-vélaire	Glottal
Durée moyenne en ms (écart type)	329 (83)	124 (60)	131 (56)	128 (58)	155 (62)

Tableau 60 : Durée moyenne (en ms) des voyelles tchèques produites en isolation et en contextes p, t, k, h (4 répétitions*3 positions*20 locutrices). L'écart type est entre parenthèse

Comme le montre le Tableau 60, les voyelles insérées en contexte glottal sont plus longues (en moyenne de 155 ms) que celles prononcées dans les autres contextes (p, t, k). Tout comme en français, les voyelles produites en isolation sont les plus longues : elles durent en moyenne 329 ms. Notons que l'écart type qui peut correspondre à presque 50% de la valeur moyenne est dû à la grande différence de durée entre les voyelles brèves et longues. Ainsi, les variations des formants des voyelles en contexte par rapport à la cible peuvent de nouveau s'expliquer par une durée vocalique plus courte.

9.3.7 Coarticulation en tchèque : Conclusion partielle

Le premier formant varie moins que le deuxième formant, comme en français. Indépendamment du contexte consonantique, les changements importants (supérieurs à 50 Hz), affectant les voyelles de moyenne et grande aperture [ɛ, a] correspondent toujours à une baisse de valeurs de F1 par rapport aux valeurs prototypiques des voyelles isolées (la cible).

Les variations du deuxième formant sont systématiques (le F2 de la plupart des voyelles tchèques dépend essentiellement de la cavité antérieure) et plus ou moins importantes selon le contexte. Le F2 subit des changements moindres en contexte glottal (notamment celui des voyelles [i:, ɛ, a, a:]) et en contexte labial (notamment le F2 de [a, o, o:]).

Le troisième formant des voyelles antérieures en contexte par rapport à celui des voyelles isolées, baisse généralement ou ne rencontre pas de variations importantes.

Enfin les variations du quatrième formant étudié pour les voyelles antérieures fermées [ɪ, i:] sont minimales.

Le Tableau 61 récapitule les mouvements par rapport à la cible des deuxième et troisième formants dus aux contextes labial, dental, palato-vélaire et glottal pour les voyelles antérieures longues et brèves, postérieures longues et brèves et les voyelles [a] et [a:]. La flèche vers le haut indique que la valeur du formant augmente, la flèche vers le bas indique qu'elle baisse. Le nombre de flèches illustre l'ampleur des mouvements formantiques. Une flèche est égale à environ 100 Hz, une flèche entre parenthèses signifie un mouvement entre 70-99 Hz et le symbole + montre un mouvement plus important au niveau des paires de voyelles brève et longue.

Formants		F2				F3			
Voyelle/contexte		p	t	k	h	p	t	k	h
Antérieures longues	i:	--	--	--	--	--	↓	--	↓
	ɛ:	↓	↓	↑↑	--	--	--	--	--
Antérieures brèves	ɪ	↓↓(↓)	↓↓(↓)	--	--	↓(↓)	↓	--	--
	ɛ	↓(↓)	↓	↑↑	--	--	--	(↓)	--
Postérieures longues	u:	--	↑↑	(↑)	(↑)				
	o:	--	↑	(↑)	(↑)				
Postérieures brèves	u	--	↑↑↑	--	(↑)				
	o	--	↑+	--	--				
[a]/ [a:]	a:	--	↑	↑(↑)	--	--	--	↓	--
	a	--	↑	↑(↑+)	--	--	↑	↓↓	(↑)

Tableau 61 : Les mouvements de F2 et F3 en fonction du groupe vocalique (voyelles antérieures/ postérieures, longues/brèves et [a]) et de la consonne subséquente (p, t, k, h). ↑ signifie une augmentation de la valeur et ↓ indique sa baisse

La lecture du Tableau 61 rend compte des effets de coarticulation en tchèque. Le contexte labial provoque par rapport à la cible :

- une baisse du F2 des voyelles antérieures, mis à part le [i:] qui est résistant à la coarticulation. Le F2 de [a/a:] n'est pas impacté.

Le contexte dental engendre :

- une baisse du F2 des voyelles antérieures, mis à part le [i:] qui est résistant à la coarticulation
- une élévation du F2 des voyelles postérieures et du [a] et [a:]

Le contexte palato-vélaire occasionne ensuite :

- une élévation du F2 et une baisse du F3 des voyelles de grande aperture [a/a:] et [ε/ε:], créant ainsi la « pince vélaire » que l'on peut voir sur le spectrogramme de [ε, ε:, a, a:], prononcés par une Tchèque native (Figure 85). Cette convergence acoustique est due à la même fréquence de résonance de la cavité antérieure et postérieure.

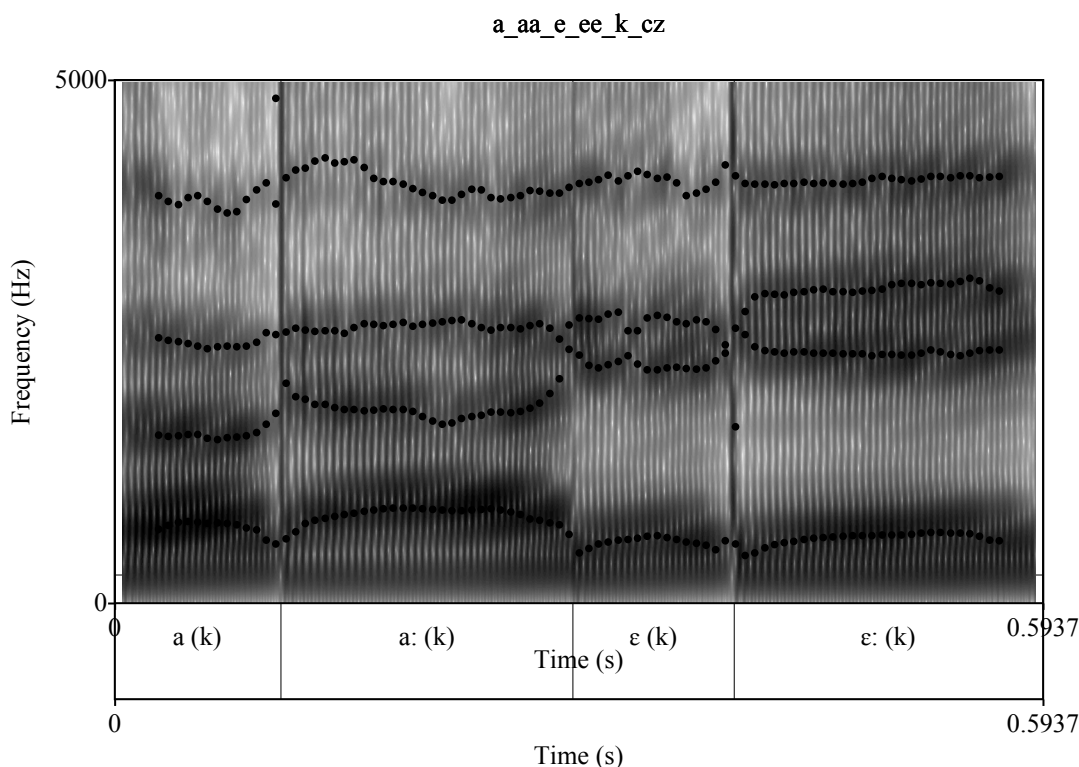


Figure 85 : Les voyelles [a, a:, ε, ε:] du tchèque en contexte kV prononcées par une Tchèque native. Notons un rapprochement des formants F2 et F3 sous forme d'une « pince vélaire »

Enfin, le contexte glottal provoque de faibles variations des formants par rapport à la cible, comme attendu: Stevens and House (1963) considèrent la glottale [h] comme un contexte neutre à cause de l'absence de constriction due à la consonne. Notons cependant que le F3 du [i:] long baisse.

Ainsi, les sons du tchèque *compatibles* - dans le sens du terme « compatibilité », utilisé par van Bergem (1993) - sont élucidés dans le Tableau 62.

Consonne	Compatibilité	Voyelle(s)
h	Toutes les voyelles, mis à part le [i:]	a, ε, i, o, u a:, ε:, o:, u:
p	arrondies	u, o
k	fermées	i, u

Tableau 62 : Contextes consonantiques qui conservent les valeurs formantiques cibles de voyelles subséquentes du tchèque

De manière générale, les voyelles longues sont plus résistantes à la coarticulation que les voyelles brèves, le [i:] étant généralement le plus résistant. Ce résultat est en accord avec la littérature (Lindblom, 1963). Le contexte engendrant les variations acoustiques les plus amples des voyelles du tchèque est le contexte dental, comme en français.

9.4 Différences entre les effets de la coarticulation en français et en tchèque

Les phénomènes de coarticulation suivent à peu près les mêmes tendances dans les deux langues : les variations du premier formant dues au contexte phonétique sont globalement moins importantes que celles du deuxième formant (ou troisième pour les voyelles antérieures dont le F3 est essentiellement affilié à la cavité antérieure), ce qui est en accord avec la littérature (Hillenbrand *et al.*, 2001; Stevens and House, 1963). Seul le contexte uvulaire provoque une forte élévation du F1, comme attendu (Stevens, 1998).

Le sens des mouvements formantiques des voyelles en contexte consonantique par rapport à la cible est souvent constant et correspond dans la majorité des cas à la modélisation par VTDemo (résultats exposés en 9.1). Le contexte labial provoque alors dans les deux langues une baisse des formants de voyelles antérieures étirées affiliés à la cavité antérieure. Le contexte dental engendre une baisse des formants dépendants de la cavité antérieure dans le cas des voyelles antérieures étirées et une élévation de ceux des voyelles postérieures. Enfin, le contexte palato-vélaire provoque une élévation du F2 et une baisse du F3 des voyelles moyennes pour lesquelles les deux formants sont affiliés aux deux cavités antérieure et postérieure qui résonnent alors à la même fréquence.

Il existe toutefois des différences entre les effets de coarticulation dans les deux langues qui sont illustrées à la Figure 86. Il s'agit des triangles vocaliques F1/F2 (en Bark) des voyelles du français (triangle du haut) et du tchèque (triangle du bas) en isolation (trait plein) et en contextes symétriques (trait pointillé) [p, t, k], puis [ɣ] ou [ɦ]. Les valeurs formantiques moyennes sont calculées à partir de productions de voyelles insérées dans des logatomes trisyllabiques, répétés quatre fois par dix locutrices françaises natives et 20 locutrices tchèques natives (valeur relevée à la moitié de la durée vocalique).

9. Propriétés acoustiques des voyelles en contextes phonétiques p, t, k, R/h : Effets de la coarticulation

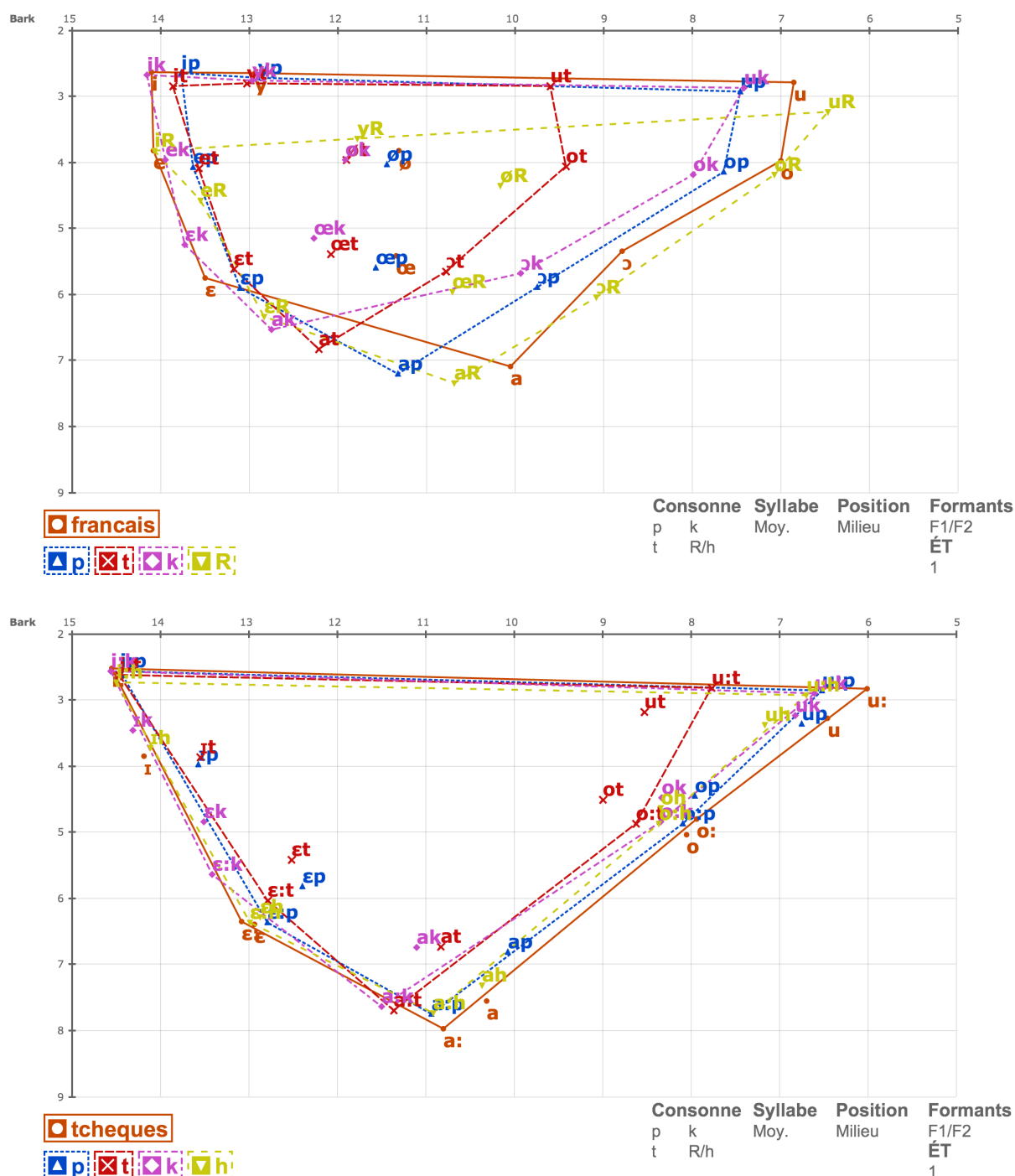


Figure 86 : Triangles vocaliques F1/F2 (en Bark) des voyelles du français (image du haut) et des voyelles du tchèque (image du bas) en contextes nul (un rond, trait plein), labial (un triangle), dental (une croix), palato-vélaire (un losange) et uvulaire ou glottal (un triangle). Moyennes calculées à partir de productions de 10 Françaises et 20 Tchèques*4 répétitions*3 syllabes (valeurs prises à la moitié de la durée vocalique)

La comparaison des triangles F1/F2 (en Bark) des voyelles du français (image du haut) et des voyelles du tchèque (image du bas) de la Figure 86 en contextes nul (marquées par un rond), labial (marquées par un triangle), dental (marquées par une croix) et palato-vélaire (marquées par un losange), met en évidence les différents effets de coarticulation qui sont les suivants :

1. Le sens des mouvements formantiques des voyelles postérieures en contexte labial. En tchèque, le formant vocalique F2 de ces dernières évolue minimalement par rapport à la cible, mis à part le [u/u:] qui augmente légèrement (de 43 Hz). Ce résultat est en accord avec (Stevens and House, 1963) étudiant l'anglais américain, qui observent peu de changements du F2 des

voyelles postérieures en contexte labial par rapport à la cible, à l'exception du [u] dont le F2 augmente légèrement (pour plus de détail sur leur étude, consulter la section 1.2). En français, le deuxième formant des voyelles postérieures [u, o, ɔ] s'élève en contexte labial par rapport à la cible de 82 Hz, 100 Hz, 170 Hz. Cette élévation ne peut donc pas s'expliquer par le locus bas de la consonne labiale mais plutôt par la centralisation acoustique attirant le F2 vers 1500 Hz (qui est le deuxième formant de la voyelle neutre modélisée par Fant (1960).

2. L'ampleur des mouvements formantiques :

- a. des voyelles antérieures. Alors que l'espace acoustique des voyelles antérieures comporte deux séries de voyelles, labiales et étirées, en français, celui du tchèque ne comporte que les voyelles étirées. Ainsi, le degré de coarticulation des voyelles [ɪ, ɛ] du tchèque est plus élevé que celui des voyelles françaises acoustiquement proches [e, ɛ]. La différence essentielle concerne la voyelle [ɛ] en contexte palato-vélaire [k] : alors que le deuxième formant de cette dernière augmente en moyenne de 88 Hz par rapport à la valeur cible en français, il augmente de 219 Hz en tchèque. Le spectrogramme de la Figure 87 illustre cette différence en mettant en contraste le [ɛ] français, prononcé en isolation et en contexte palato-vélaire par une Française native et le [ɛ] tchèque, prononcé dans ces mêmes environnements phonétiques par une Tchèque native. La différence de l'ampleur de coarticulation des voyelles antérieures entre les deux langues n'est néanmoins pas aussi importante que celle des voyelles postérieures, décrite dans le paragraphe qui suit.

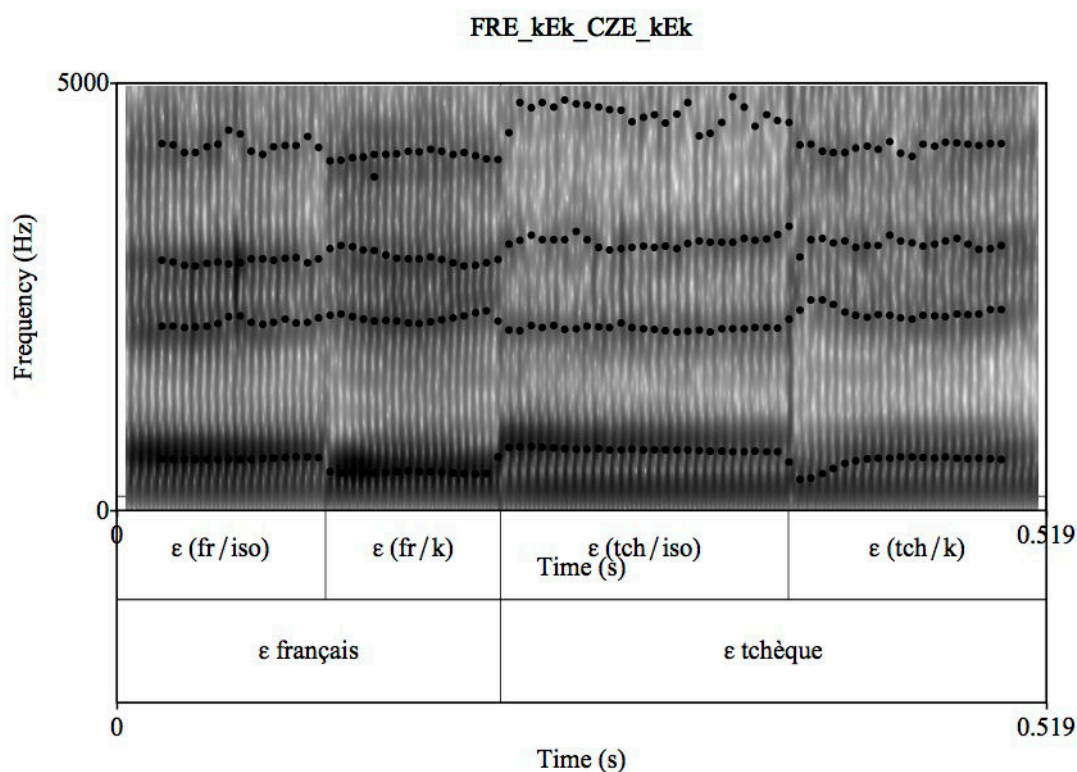


Figure 87 : Spectrogramme de la voyelle [ɛ] prononcée en isolation et en contexte kVk par une Française parisienne et une Tchèque de Bohême. Le F2 du [ɛ] en contexte kVk augmente par rapport au [ɛ] isolé davantage en tchèque qu'en français

Ce résultat est en concordance avec l'hypothèse de Manuel and Krakow (1984) qui expliquent qu'une plus grande variation acoustique des voyelles dans un système vocalique plus restreint ne menace pas les contrastes vocaliques.

- b. des voyelles postérieures. L'espace acoustique des voyelles postérieures ne comporte qu'une série de voyelles dans les deux langues. Une plus grande variation du F2 (mais non pas du F1) ne peut donc pas menacer la neutralisation des contrastes, comme ce serait le cas des voyelles antérieures du français. L'élévation du F2 de ces voyelles en

contextes dental et palato-vélaire est donc grande, notamment en français. Ainsi, le deuxième formant des voyelles du français /u, o, ɔ/ augmente en contexte dental par rapport à la cible de 440 Hz, 393 Hz et 389 Hz respectivement alors que celui des voyelles tchèques /u, o/ augmente dans le même contexte par rapport à la cible, respectivement de 310 Hz et de 162 Hz. En contexte palato-vélaire, le deuxième formant de la série française /u, o, ɔ/ augmente de 76 Hz, 153 Hz et 207 Hz respectivement alors que celui de la série tchèque /u, o/ n'augmente par rapport à la cible que de 52 Hz et 48 Hz respectivement. La grande élévation du F2 des voyelles postérieures en contexte impacte également leur caractère focal, qui est plus affaibli en français qu'en tchèque. Alors qu'en tchèque, les voyelles postérieures [u, u:, o, o:] gardent une distance F2-F1 inférieure à 500 Hz pratiquement dans tous les contextes, mis à part le [o] en contextes dental et palato-vélaire, cette distance pour les voyelles françaises [u, o, ɔ] est supérieure à 500 Hz dans pratiquement tous les contextes, à l'exception du [o] en contexte uvulaire et labial.

- c. de [a]. Du fait de sa réalisation pharyngale en position isolée, il s'agit de la voyelle qui change maximalement son patron formantique selon le contexte dans lequel elle apparaît. L'ampleur de coarticulation de [a] avec les consonnes environnantes est plus importante en français qu'en tchèque, comme nous pouvons l'observer à la Figure 88. Ce graphe met en contraste l'évolution des formants moyens F1, F2, F3 (en Bark) de la voyelle [a] prononcée dans des logatomes trisyllabique CVCVCVC où C pour le français correspond aux consonnes [p, t, k, ʁ] (marquée avec une croix) et pour le tchèque aux consonnes [p, t, k, h] (marquée avec un rond). Les valeurs formantiques des voyelles isolées servent de référence et correspondent aux traits horizontaux qui traversent l'image. La voyelle est postérieure dans les deux langues quand elle est prononcée en isolation. En français, le F2 de [a] passe de 1300 Hz en isolation à 1571 Hz en contexte labial, à 1808 Hz en contexte dental et à 1976 Hz en contexte palato-vélaire. En tchèque, le F2 de [a] passe de 1344 Hz en isolation à 1457 Hz en contexte dental et à 1520 Hz en contexte palato-vélaire.

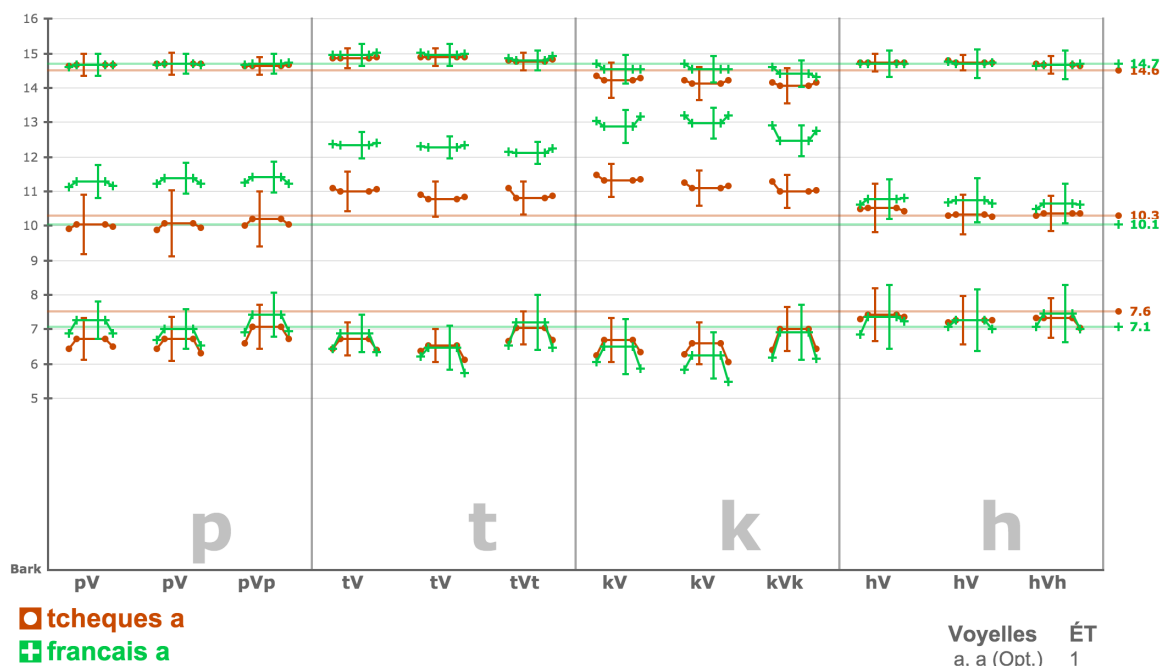


Figure 88 : Les formants moyens F1, F2 et F3 (en Bark) de la voyelle [a] du français (représenté avec une croix) et du tchèque (représenté avec un rond) en isolation (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ] pour le français et [p, t, k, h] pour le tchèque. Les moyennes sont calculées à partir de productions de 10 Françaises/20 Tchèques* 4 répétitions

Les différents effets de coarticulation élucidés ci-dessus affectent ainsi la similarité acoustique des voyelles du français et du tchèque en contexte. Ce point fera l'objet de la section suivante.

9.5 Similarité acoustique des voyelles du français et du tchèque en contextes symétriques p, t, k

Puisque le français et le tchèque diffèrent notamment par l'ampleur de la coarticulation, comme nous l'avons montré dans la section précédente, la similarité acoustique des voyelles des deux langues varie en fonction du contexte phonétique et elle ne peut pas être prédite à partir de la similarité acoustique des voyelles isolées.

Alors que le chapitre 8 a montré qu'aucune paire des voyelles isolées du français et du tchèque n'est *identique* au niveau de tous les formants comparés (résultats du test-t se trouvant à la page 44 des annexes), qu'en est-il des voyelles prononcées en contextes labial, dental et palato-vélaire ?

Nous émettons l'hypothèse qu'il existe des voyelles du français acoustiquement *identiques* aux voyelles tchèques prononcées en contextes symétriques labial (p), dental (t) et palato-vélaire (k). Autrement dit, nous stipulons que certaines voyelles du français et du tchèque insérées dans les mêmes contextes consonantiques ne diffèrent pas par la valeur des formants comparés (F1, F2 des voyelles postérieures, F1-F3 des voyelles antérieures, F1-F4 du [i]).

L'hypothèse nulle consiste à dire qu'il n'existe pas de voyelles du français acoustiquement identiques aux voyelles du tchèques lors de leur production dans les mêmes contextes phonétiques.

Afin de vérifier notre hypothèse, nous avons conduit un test-t bilatéral pour une comparaison des moyennes formantiques des voyelles du français et du tchèque en contextes [p, t, k]. Les paires examinées étaient de nouveau [i]fr - [i:]tch, [e]fr - [ɪ]tch, [ɛ]fr - [ɛ]tch, [ɛ]fr - [ɛ:]tch, [a]fr - [a]tch, [o]fr - [o]tch, [o]fr - [u]tch, [ɔ]fr - [o]tch, [ɔ]fr - [a]tch, [u]fr - [u:]tch, [œ]fr - [ɛ]tch. Les voyelles postérieures ont été comparées par la valeur de leur F1-F2, les voyelles antérieures par la valeur des F1 à F3 et la paire [i, i:] par la valeur des F1 à F4. Le résultat complet se trouve en annexe du chapitre 9, aux pages 73 - 75 des annexes.

Malgré des valeurs moyennes très proches, le résultat du test-t montre qu'il s'agit de valeurs souvent significativement différentes ce qui s'explique par l'étendue des valeurs qui varie selon les langues (différence au niveau de l'écart type, marqué entre parenthèses). La valeur du test-t bilatéral est en effet $p < 0,05$ pour les formants de toutes les voyelles comparées, mis à part le F1 de la paire [ɛ-ɛ] en contexte pVp où $p = 0,07$, de la paire [a-a] en contexte tVt où $p = 0,12$ et de la paire [u-u:] en contexte tVt où $p = 0,36$ et en contexte kVk où $p = 0,41$, puis mis à part le F2 de la paire [e-i] en contexte tVt où $p = 0,46$ et pour le F3 de la paire [i-i:] en contexte pVp où $p = 0,35$ et de la paire [a-a] en contexte pVp où $p = 0,63$. Cette différence d'étendue de valeurs a déjà été soulevée par (Flege, 2003) qui a observé que les voyelles de deux langues différentes sont rarement *identiques*.

Les tchécophones n'ont de nouveau jamais affaire à des voyelles acoustiquement *identiques* dans l'apprentissage des voyelles du français en contextes p, t, k car il existe des différences significatives entre les moyennes d'un ou de plusieurs formants comparés. Notre hypothèse de départ est alors rejetée et l'hypothèse nulle est retenue.

De même que dans le cas des voyelles isolées, nous utiliserons ici le terme "hautement similaire" au lieu de *identique* pour décrire les voyelles acoustiquement proches dont les formants se trouvent à un écart type ou moins des formants moyens des voyelles tchèques. Les voyelles du français *similaires* aux voyelles du tchèque sont de nouveau celles dont les formants moyens se trouvent entre un et deux écarts types. Enfin les voyelles du français acoustiquement *nouvelles* se réalisent avec un ou plusieurs formants à plus de deux écarts types de ceux des voyelles du tchèques.

La similarité acoustique des voyelles des deux langues est élucidée pour chaque contexte séparément. Les tableaux ainsi que les triangles vocaliques et les spectrogrammes montrant la similarité acoustique des voyelles du français et du tchèque en contextes labial, dental et palato-vélaire se trouvent en annexe du chapitre 9 (pages 75 – 83 des annexes). Le Tableau 63 résume cette similarité en termes de nombre d'écarts types des formants moyens F1, F2 (toutes les voyelles) et F1, F2 et F3 (voyelles antérieures) et rappelle pour comparaison la similarité des voyelles isolées.

Voyelles FR	Similarité +++ (≤ 1 ET)	Similarité ++ (1 - 1,5 ET)	Similarité + (1,5 – 2 ET)	Similarité 0 (> 2 ET)
i		i: (p, t, k)	i: (0)	
e	ɪ (p, t)	ɪ (k)	ɪ (0)	
ɛ	ɛ: (p, t, k)		ɛ (0)	
a	a (0) a: (p)		a:ɛ: (t) ɛ: (k)	
u		u/u: (0, p, t, k)		
o	o (p) o/o: (k)	o (t)	u (0)	
ɔ		o (0)	a (p)	(t, k)
y				(0, p, t, k)
ø				(0, p, t, k)
œ	ɛ (t)		ɛ (p), ɛ: (k)	(0)

Tableau 63 : Similarité des voyelles du français (colonne à gauche) et du tchèque en contextes [p, t, k] et nul (0), précisé entre parenthèses. – [= 1 ET = Similarité +++, 1 – 1,5 ET = Similarité ++, 1,5 – 2 ET = Similarité +, 2 et + ET = Similarité 0

La première colonne du Tableau 63 représente les dix voyelles du français (FR) et les colonnes suivantes montrent de quelle(s) voyelle(s) du tchèque se rapproche le plus chaque voyelle du français en précisant le(s) contexte(s) (indiqué(s) entre parenthèses).

Le tableau met en évidence plusieurs phénomènes :

1. Aucune des dix voyelles du français n'est « hautement similaire » dans tous les contextes étudiés. La voyelle [e] est « hautement similaire » au [ɪ] tchèque en contextes labial et dental. La voyelle [ɛ] est « hautement similaire » en contextes [p, t, k] au [ɛ:] long tchèque. Les autres voyelles qui peuvent être “hautement similaires”, selon le contexte, sont le [a, o, œ]. Remarquons que le [a] français est « hautement similaire » en contextes nul et labial au [a/a:] tchèque et *similaire* en contextes dental et palato-vélaire au [a:] et davantage au [ɛ:] long tchèque.
2. Il existe une similarité acoustique entre des voyelles retranscrites avec des symboles phonétiques différents. Ainsi, selon le contexte, le [e] français est acoustiquement proche du [ɪ] tchèque. Le [a] français est quant à lui proche, selon le contexte, du [a/a:] et [ɛ:] tchèque. Le [o] du français se rapproche acoustiquement du [o/o:] ou du [u] tchèque et enfin, le [ɔ] français est

similaire au [o] ou au [a] tchèque. En effet, seules les voyelles françaises [i, ε, u] ne sont dans aucun contexte *similaires* aux sons tchèques retranscrits avec un autre symbole phonétique.

3. Les voyelles hautes [i, u] ne sont jamais « hautement similaires » du point de vue acoustique. Elles se trouvent entre un et deux écarts types de la moyenne de F1, F2, (F3) du [i:, u:] tchèque respectivement.
4. Les voyelles véritablement *nouvelles*, dans tous les contextes, sont le [y, ø]. Deux autres voyelles peuvent être *nouvelles*, selon le contexte : le [ɔ] en contextes dental et palato-vélaire, puis le [œ] en isolation. Dans les autres contextes, ces voyelles sont *similaires*. En contexte dental, le [œ] est même « hautement similaire » au [ε] tchèque (à moins d'un écart type sur le plan F1/F2 et F2/F3).
5. Ainsi, selon le contexte, la même voyelle [œ] peut être *nouvelle*, *similaire* ou « hautement similaire » au [ε:/ε], ce qui est illustré à la Figure 89, indiquant les valeurs formantiques moyennes du [œ] français, prononcé par 10 Françaises, (quatre fois dans des logatomes CVCVCVC où C correspond à l'une des consonnes [p, t, k, ʁ]) et du [ε] tchèque, prononcé par 20 Tchèques (quatre fois dans des logatomes CVCVCVC où C correspond à l'une des consonnes [p, t, k, h]). L'écart type affiché est de 1 et les valeurs sont transformées en Bark.

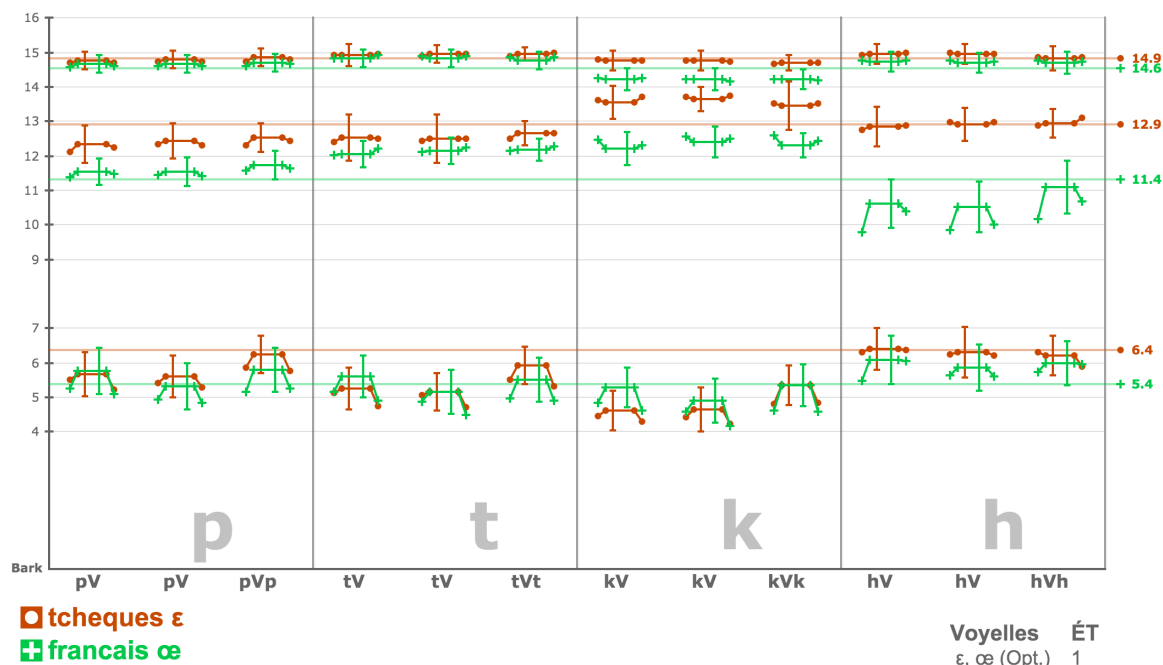


Figure 89 : Les formants moyens F1, F2, F3 (en Bark) du [œ] français (représenté avec une croix, calculés à partir de 4 répétitions par 10 Françaises) et du [ε] tchèque (représenté avec un rond, calculés à partir de 4 répétitions par 20 Tchèques) en contexte nul et dans les logatomes CVCVCVC où C = [p, t, k, R/ h]. Les traits horizontaux représentent les formants des voyelles isolées. L'écart type est de 1

La Figure 89 montre que le [œ] français est « hautement similaire » au [ε] tchèque en contexte dental, puis *similaire* en contexte labial et enfin son F2 se trouve à plus de deux écarts types du F2 de [ε] en contexte palato-vélaire. Les contextes [ʁ] pour le français et [h] pour le tchèque ne sont donnés qu'à titre indicatif.

Le chapitre 9 met ainsi en évidence que la similarité acoustique des voyelles du français et du tchèque varie fortement selon le contexte phonétique.

10. Propriétés acoustiques des voyelles initiales, médianes et finales des logatomes : Effets de la position prosodique

Dans la société des sons, il faut s'adapter pour exister et assurer ses arrières. Les sons occupent en effet dans le 'mot phonétique' des positions sociales plus ou moins avantageuses.

(Wioland, 2005, p. 67)

Résumé : Le **chapitre dix** met en évidence l'effet de la position prosodique sur les propriétés acoustiques des voyelles du français et du tchèque. Le résultat montre que le facteur de position (initiale, médiane et finale) affecte essentiellement la durée vocalique et dans une moindre mesure les valeurs formantiques. Dans les deux langues, un allongement de la voyelle finale par rapport aux voyelles initiale et médiane est observé. Les variations acoustiques concernent ensuite essentiellement le premier formant, notamment des voyelles de grande et moyenne aperture, qui baisse en position médiane en français, et en position médiane et initiale en tchèque. De cette façon, la voyelle finale – qui est systématiquement la plus longue – se réalise avec un premier formant maximal de toutes les voyelles brèves du tchèque et des voyelles mi-fermées [e, ø, o] et du [a] français. Cependant la neutralisation des contrastes entre les voyelles moyenne e/ɛ, ø/œ et o/ɔ en position non-finale n'est pas confirmée ce qui peut s'expliquer par notre choix d'un corpus basé sur des non-mots (toujours hyperarticulés selon Durand (1985)).

Dans la parole naturelle, les déplacements des formants ne peuvent pas être expliqués exclusivement par la nature des consonnes environnantes (van Bergem, 1994). Il est connu que l'ampleur des variations acoustiques d'une voyelle dépend entre autre de sa position dans le mot qui peut être forte (accentuée) ou faible (inaccentuée), notamment dans les langues à accent mobile (Garde, 1968). On observe en anglais, langue où l'accent a un rôle contrastif, une centralisation systématique des segments non accentués (Delattre, 1964) ce qui induit une réduction de type phonologique (Fourakis, 1991). Flemming (2006) précise que pour que la neutralisation des contrastes vocaliques ait lieu, il faut que la voyelle en syllabe inaccentuée soit très brève. En anglais, le schwa dans des mots insérés dans des phrases cadres ne dure en moyenne que 64 ms en position non-finale du mot lexical (Flemming and Johnson, 2007). En suédois, les voyelles en position inaccentuée sont également abrégées et les cibles ne sont pas atteintes. Ce phénomène est décrit par Lindblom (1963) en termes de 'target undershoot'. La substitution des voyelles inaccentuées par un schwa n'est néanmoins pas systématique dans cette langue (de même qu'en néerlandais par exemple) car elle ne concerne que certains mots ce qui amène van Bergem (1993) à décrire ce phénomène en termes de *réduction vocalique lexicale* plutôt que de réduction phonologique.

Le français et le tchèque sont des langues à accent fixe ayant un rôle démarcatif. Le français est défini entre autre comme une *langue sans accent* (Rossi, 1980), une *langue à accent « probabilitaire »* (Fonagy, 1980) ou encore *langue à frontières* (Vaissière, 2010). L'accent lexical sous-jacent (hérité du latin), abstrait, est sur la dernière syllabe du mot. Il est de type quantitatif et il peut ou non être réalisé dans la chaîne parlée, selon la position du mot dans l'énoncé. S'il est en fin d'un mot prosodique, l'accent se manifeste par un allongement final et des variations du fondamental sur la syllabe finale (Wioland, 2005). Il existe également une tendance en français moderne à renforcer le début de mot par

un accroissement de tension. La surtension des plis vocaux et des organes articulatoires de même que la pression sous-glottique accrue se manifestent par une intensité plus forte et un fondamental plus élevé. Un renforcement initial en français à des niveaux supérieurs au mot est décrit par exemple par Fougeron (2001). En tchèque, l'accent lexical est placé sur la syllabe initiale d'une mesure rythmique (Palková, 1997).

Dans les deux langues, un allongement final est observé (Rigault, 1970). Le phénomène d'allongement final semble universel. Lindblom (1978) précise que malgré son caractère naturel, l'allongement final est un phénomène appris et spécifique à chaque langue. Son ampleur peut donc varier d'une langue à l'autre.

Ainsi, suivant la position de la voyelle au sein du constituant prosodique, l'on peut s'attendre à des variations de la durée, des formants, de l'intensité et de la fréquence fondamentale. Nous nous intéressons ici aux changements duratifs et spectraux qui sont les deux paramètres acoustiques étudiés dans cette thèse. Nous nous posons alors les questions suivantes :

1. Quel est le rapport de durée entre les syllabes initiale, médiane et finale dans les logatomes des deux langues ?
2. L'allongement final est-il plus important en français (où il coïncide avec l'accent qualitatif lexical sous-jacent) qu'en tchèque qui a un accent initial ?
3. Les voyelles longues du tchèque en position médiane sont-elles réduites dans la même mesure que les voyelles brèves ?
4. La syllabe médiane des logatomes est-elle phonétiquement plus réduite que la syllabe finale (plus longue) et initiale (produite avec plus de force (Straka, 1964) ?
5. La réduction spectrale est-elle de la même ampleur dans les deux langues ?
6. Le contraste acoustique entre les voyelles moyennes mi-fermées et mi-ouvertes du français est-il neutralisé en position inaccentuée des logatomes ?

Dans l'étude de l'influence de la position prosodique sur les valeurs formantiques et la durée des voyelles, nous avons utilisé le corpus décrit dans la partie méthodologique (section 2.3). Le corpus ainsi que le calcul des valeurs moyennes est récapitulé dans le Tableau 64.

	CORPUS
français	10 voyelles orales du français [i, e, ε, y, ø, œ, u, o, ɔ, a], insérées dans des logatomes pVpVpVp, tVtVtVt, kVkVkVk, RVRVRVR d'une phrase cadre, exemple : « Le mot papapape peut bien coller. », répétées 4 fois
tchèque	10 voyelles monophthongues du tchèque [ɪ, i:, ε, ε:, u, u:, o, o:, a, a:], insérées dans des logatomes pVpVpVp, tVtVtVt, kVkVkVk, hVhVhVh d'une phrase cadre, exemple : « Slovo papapap pŭsobí divně ¹⁸ . », répétées 4 fois
	LOCUTRICES
françaises	10 Françaises natives non-méridionales ayant grandi ou longtemps séjourné en Ile-de-France
tchèques	20 Tchèques natives de la région de Bohême
	MESURES
Calcul de la moyenne formantique	A partir de la valeur relevée à la moitié de la durée vocalique (10/ 20 locutrices*4 répétitions*4 contextes consonantiques confondus). Nous avons ainsi obtenu les formants moyens des voyelles initiale CVCVCVC, médiane CVCVCVC et finale CVCVCVC

Tableau 64 : Corpus, locutrices et mesures des moyennes formantiques des voyelles du français et du tchèque prononcées dans 3 positions du logatomes

Dans l'étude des voyelles françaises et tchèques dans trois positions du logatome différentes, nous émettons l'hypothèse que les valeurs des formants et de la durée vocalique changent significativement d'une syllabe à l'autre dans des logatomes trisyllabiques.

L'hypothèse nulle consiste à dire qu'il n'existe pas de différences significatives entre les valeurs des formants et la durée des voyelles selon leur position dans le logatome (initiale, médiane, finale).

10.1 Réduction due à la position prosodique en français

La réduction phonétique des voyelles du français engendrée par la position prosodique a été étudiée par de nombreux chercheurs.

Fougeron (2001) étudiant les voyelles [i] et [ã] prononcées par quatre locuteurs Français parisiens (un homme et trois femmes) à l'initial des constituants prosodiques (syllabe, mot lexical, groupe accentuel et intonatif) n'a pas observé de variations systématiques de la durée vocalique en fonction du type de constituant.

A l'inverse, Gendrot and Gerdes (2009) ont montré, à partir de 30 heures de la parole journalistique, qu'il existe en français un allongement des voyelles se trouvant aux frontières des constituants prosodiques (mots lexicaux, groupes accentuels ou intonatifs). En effet, les voyelles

¹⁸ Nous traduisons « Le mot papapap donne une impression bizarre. »

initiales s'avèrent plus longues que les voyelles à l'intérieur des groupes prosodiques. Cependant les voyelles en position finale des groupes intonatifs sont les plus longues (leur durée par rapport aux voyelles à l'intérieur du constituant double et parfois triple), comme attendu.

La réalisation spectrale des voyelles dépend de leur position prosodique. Selon Gendrot et Gerdes (2009), l'espace acoustique est minimal pour les voyelles se situant à l'intérieur d'un constituant prosodique alors que les formants des voyelles occupant les positions initiale ou finale atteignent des valeurs plus extrêmes. Les auteurs notent par ailleurs que plus le constituant prosodique est important (mot lexical < mot prosodique < groupe intonatif), plus l'espace vocalique s'agrandit.

Selon Wioland (2005), les voyelles du français sont centralisées en syllabe non-finale d'un mot prosodique ou lexical. Ainsi les voyelles inaccentuées fermées [i, y, u] s'ouvrent, la voyelle ouverte [a] se ferme et l'opposition entre les voyelles mi-fermées et mi-ouvertes est neutralisée car l'aperture devient moyenne, notamment en syllabe inaccentuée CV. Selon Delattre (1969), les voyelles mi-fermées [e, o] ne subissent pas de centralisation spectrale (avec une élévation de F1) en position inaccentuée.

10.1.1 Variation de la durée suivant la position prosodique en français

Dans notre étude, nous nous sommes intéressée à la variation temporelle des voyelles suivant les positions initiale, médiane et finale au sein du logatome. La durée moyenne (toutes les voyelles et contextes confondus) de la syllabe finale en français est de 133 ms (ET de 53 ms) alors que celle de la syllabe initiale et médiane est de 92 ms (ET de 45 ms) et 93 ms (ET de 35 ms) respectivement. Le rapport des voyelles en positions initiale et finale, de même qu'en positions médiane et finale est de 1 : 1,4.

Afin de mesurer l'effet de la position dans le logatome sur la durée de la voyelle, nous avons effectué une analyse ANOVA à un facteur qui est la « Syllabe » et qui comprend trois modalités : initiale, médiane et finale. La valeur ANOVA $F(2, 4790) = 431$ avec $p < 0,05$, comme indiquée dans le Tableau 65, montre que l'effet de la position sur la durée vocalique est significatif.

Source	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	2	1741712,780	870856,390	430,758	< 0,0001
Erreur	4790	9683864,657	2021,684		

Tableau 65 : Résultat de l'ANOVA à un facteur mesurant l'effet de la syllabe sur la durée vocalique

Pour vérifier à quelles positions est dû l'effet général sur la durée vocalique, nous avons conduit un test a posteriori de Fisher dont le résultat est indiqué dans le Tableau 66.

Contraste	Différence	Différence standardisée	Valeur critique	Pr > Diff
FIN vs BEG	40,658	25,556	1,960	< 0,0001
FIN vs MID	40,223	25,283	1,960	< 0,0001
MID vs BEG	0,434	0,273	1,960	0,785

Tableau 66 : PLSD de Fisher pour le facteur « Syllabe » avec niveau de significativité de 5 %, FIN = syllabe finale, MID = syllabe médiane, BEG = syllabe initiale

Le test Fisher indique que l'effet global de la position dans le mot sur la durée vocalique est principalement dû à des différences significatives de durée entre la syllabe finale et les syllabes initiale ou médiane. En revanche, la différence durée des syllabes initiale et médiane n'est pas significative. Le graphe de la Figure 90 montre qu'en effet, la voyelle finale est plus longue que la voyelle initiale ou médiane.

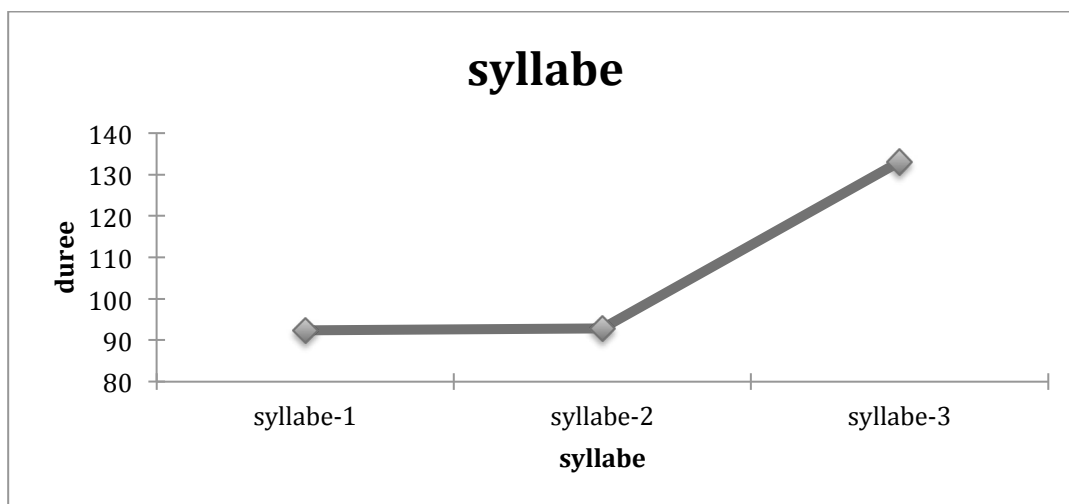


Figure 90 : Variation de la durée vocalique selon la position dans le mot : Syllabe 1 = initiale, Syllabe 2 = médiane, Syllabe 3 = finale

Notre hypothèse personnelle, à savoir qu'il existe en français une différence de durée vocalique selon la position de la voyelle dans le mot, est confirmée et l'hypothèse nulle peut être rejetée. Ce résultat est en accord avec l'allongement prépausal reporté par Gendrot and Gerdes (2009) pour les voyelles du français et avec l'étude de Fougeron (2001) indiquant que les voyelles à l'initial d'un constituant prosodique ne sont pas nécessairement allongées en français.

10.1.2 Variation des formants suivant la position prosodique en français

Pour mesurer l'effet de la position dans le logatome sur les formants vocaliques en français, nous avons conduit une analyse ANOVA à deux facteurs, voyelles et syllabe, pour chaque formant. Le facteur « Voyelles » est à dix modalités et correspond aux voyelles [i, e, ε, y, ø, œ, a, u, o, ɔ] et le facteur « Syllabe » à trois modalités correspond aux syllabes initiale, médiane et finale.

Les valeurs moyennes des formants avec les écarts types entre parenthèses des voyelles du français selon la position initiale, médiane et finale se trouvent dans le Tableau 67 de la page 204. Nous y remarquons que le premier formant vocalique de la voyelle médiane est systématiquement plus bas que celui de la voyelle initiale et finale et que cette baisse est plus grande dans le cas des voyelles mi-ouvertes [ε, œ, ɔ] et de la voyelle ouverte [a].

Formants Voyelles/ Syllabes		F1			F2			F3		
		Initiale	Médiane	Finale	Initiale	Médiane	Finale	Initiale	Médiane	Finale
i	fermées	308 (52)	302 (64)	319 (72)	2459 (242)	2454 (238)	2448 (251)	3474 (253)	3467 (263)	3494 (256)
y		314 (51)	300 (58)	311 (56)	1958 (215)	1952 (247)	1975 (205)	2600 (212)	2619 (217)	2589 (183)
u		311 (34)	304 (38)	307 (37)	930 (203)	901 (233)	895 (202)			
e	mi-fermées	439 (53)	428 (52)	452 (59)	2317 (185)	2322 (193)	2328 (187)	3055 (160)	3074 (163)	3053 (170)
ø		422 (52)	419 (51)	443 (54)	1569 (212)	1592 (237)	1628 (191)	2670 (203)	2692 (220)	2645 (172)
o		437 (52)	429 (47)	441 (48)	936 (182)	953 (189)	957 (156)			
ɛ	mi-ouvertes	650 (89)	604 (93)	634 (75)	2133 (170)	2141 (194)	2145 (168)	2962 (163)	2974 (172)	2951 (164)
œ		615 (78)	570 (88)	612 (72)	1639 (193)	1655 (217)	1693 (159)	2753 (195)	2745 (201)	2744 (194)
ɔ		653 (75)	610 (79)	639 (62)	1246 (162)	1265 (176)	1282 (146)			
a	ouverte	787 (103)	753 (106)	822 (109)	1707 (260)	1717 (268)	1658 (210)	2809 (217)	2815 (216)	2770 (204)

Tableau 67 : Valeurs moyennes de F1 et F2 (toutes les voyelles), F1 à F3 (voyelles antérieures) des voyelles du français, prononcées en syllabes initiale, médiane et finale dans des logatomes CVCVCVC où C = [p, t, k, ʁ] (valeur prise à la moitié de la durée* 4 répétitions* 4 contextes confondus* 10 locutrices). L'écart type est entre parenthèses. Les variations selon la syllabe les plus amples sont grisées

Le résultat du test ANOVA avec $F(29, 4763) = 890$ avec $p < 0,05$ (voir Tableau 68) indique que l'effet global de la position dans le logatome sur le premier formant des dix voyelles du tchèque est significatif.

F1	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	29	120289099,619	4147899,987	889,920	< 0,0001
Erreur	4763	22200240,236	4660,978		

Tableau 68 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur le F1 des voyelles du français

Comme indiqué dans le Tableau 69, l'effet du premier facteur « Voyelle » est grand, car $F(9, 4763) = 2842$, $p < 0,05$. La valeur ANOVA du deuxième facteur « Syllabe » est $F(2, 4763) = 66$ avec p

$< 0,05$. L'effet combiné des deux facteurs « Voyelle*Syllabe » est significatif et il est exprimé par $F(18, 2763) = 5$ avec $p < 0,05$.

	F1	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle		9	119231499,924	13247944,436	2842,310	$< 0,0001$
syllabe		2	615260,473	307630,236	66,001	$< 0,0001$
voyelle*syllabe		18	442339,222	24574,401	5,272	$< 0,0001$

Tableau 69 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur le F1 des voyelles du français

Le test a posteriori de Fisher montre à quels groupes est dû l'effet combiné des deux facteurs et le résultat complet se trouve dans les annexes à la page 69. Il montre que le premier formant des voyelles [i, e, ø, ε, œ, ɔ, a] est alors significativement plus bas dans la syllabe médiane par rapport à la syllabe finale et/ou initiale. La réduction vocalique des voyelles (mi) - ouvertes [ε, ɔ, œ, a] en syllabe médiane se traduisant par un premier formant plus bas, correspond aux attentes à propos de la centralisation des voyelles en position inaccentuée. En revanche la baisse de F1 du [i] fermé et des [e, ø] mi-fermés en syllabe médiane, corrélée à une aperture plus petite, n'est pas en concordance avec la littérature (Wioland, 2005). Notons que le F1 des voyelles (mi-)fermées [y, u, o] ne varie pas selon la position dans le logatome comme le montre le résultat du test Fisher. La Figure ci-dessous indique la variation du premier formant des voyelles orales du français due à la position dans le mot : initiale (BEG, en rouge), médiane (MID, en vert) et finale (FIN, en bleu).

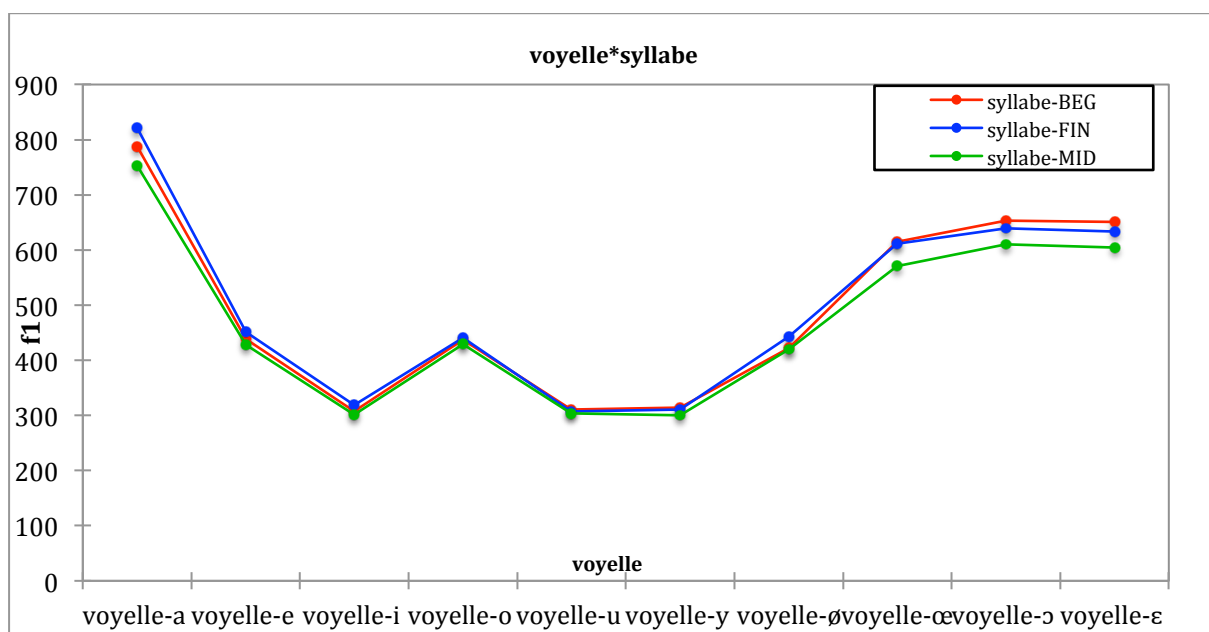


Figure 91 : F1 des voyelles orales du français en position initiale (en rouge), médiane (en vert) et finale (en bleu), calculé à partir de la valeur centrale*4 répétitions*4 contextes confondus*10 locutrices

Le Tableau 70 indique la valeur ANOVA $F(29, 4763) = 1021$ avec $p < 0,05$ qui montre l'effet global de la position de la syllabe sur le deuxième formant vocalique.

F2	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	29	1251580814,220	43157959,111	1020,618	< 0,0001
Erreur	4763	201408690,038	42286,099		

Tableau 70 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur le F2 des voyelles du français

Comme attendu, l'effet de la variable nominale « Voyelle » sur les formants est grand, indiquant que la valeur du F2 varie significativement selon la voyelle. La valeur ANOVA pour ce facteur, comme indiqué dans le tableau, est $F(9, 4763) = 3286$ avec $p < 0,05$. En revanche, la valeur ANOVA montre que l'effet du facteur « Syllabe » sur le F2 (avec $F(2, 4763) = 1$ avec $p = 0,291$) et l'effet combiné des facteurs « Voyelle*Syllabe » (avec $F(18, 4763) = 1$ avec $p = 0,115$) n'est pas significatif (voir Tableau 71).

F2	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle	9	1250401577,380	138933508,598	3285,560	< 0,0001
syllabe	2	104492,394	52246,197	1,236	0,291
voyelle*syllabe	18	1074744,445	59708,025	1,412	0,115

Tableau 71 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur le F2 des voyelles du français

La Figure 92 illustre la variation du deuxième formant des dix voyelles orales du français en fonction de la position dans le logatome qui est soit initiale, soit médiane, soit finale. Nous y remarquons une petite variation (d'environ 50 Hz) de la valeur du deuxième formant suivant la position dans le logatome pour les voyelles moyennes [ø] et [œ].

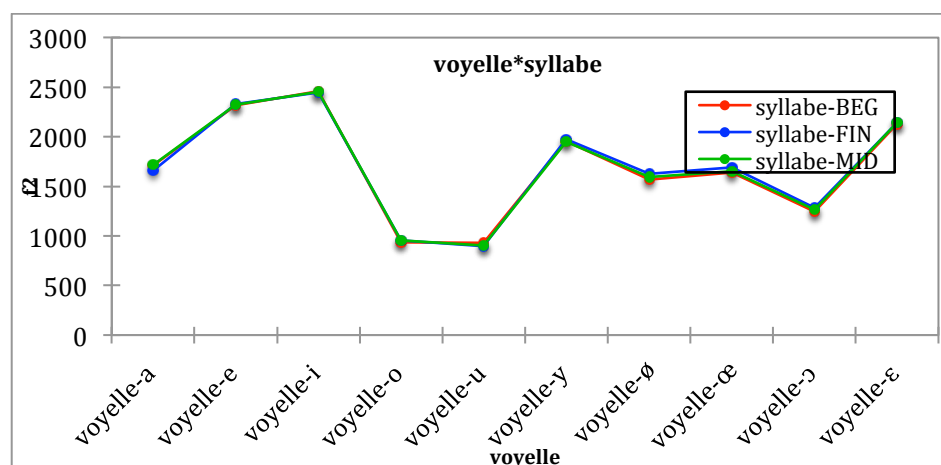


Figure 92 : F2 des voyelles orales du français en positions initiale (en rouge), médiane (en vert) et finale (en bleu), calculé à partir de la valeur prise à la moitié de la durée vocalique* 4 répétitions* 4 contextes confondus* 10 locutrices

Comme l'indique le Tableau 67 de la page 204, le F2 de [ø] est de 1628 Hz et le F2 de [œ] est de 1693 Hz en syllabe finale fermée. En syllabe initiale ouverte, la valeur du F2 est plus basse à savoir de 1569 Hz pour [ø] et de 1639 Hz pour [œ]. Cette variation pourrait être expliquée par un autre facteur que la position dans le logatome et il s'agirait du type de la syllabe, fermée ou ouverte, qui déterminerait la réalisation de la voyelle moyenne. Le respect de la loi de position a pu en effet avoir un impact (quoique peu important) sur la valeur du deuxième formant de la paire ø/œ qui est plus élevée en syllabe fermée qu'en syllabe ouverte. Il est à noter que malgré cette variation du F2, les

timbres mi-fermés et mi-ouverts restent distingués au niveau du premier formant dans toutes les positions du logatome et cela dans les trois paires des voyelles moyennes (pour les valeurs moyennes exactes, se référer au Tableau 67).

Enfin, au niveau du troisième formant, l'effet global de la syllabe est donné par la valeur $F(29, 4763) = 195$ avec $p < 0,05$, comme le montre le Tableau 72.

	F3	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle		29	272469422,711	9395497,335	195,645	< 0,0001
Erreur		4763	228733982,114	48023,091		

Tableau 72 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur le F3 des voyelles du français

Le Tableau 73 indique ensuite l'effet du facteur « Voyelle » sur le formant F3 qui correspond à la valeur ANOVA $F(9, 4763) = 627$ avec $p < 0,05$. La valeur ANOVA pour le deuxième facteur « Syllabe » est $F(2, 4763) = 6$ avec $p < 0,05$. Enfin, l'effet combiné des deux facteurs « Voyelle*Syllabe » est donné par $F(18, 4763) = 1$ avec $p = 0,23$ et il est non significatif.

	F3	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle		9	270795101,292	30088344,588	626,539	< 0,0001
syllabe		2	615449,590	307724,795	6,408	0,002
voyelle*syllabe		18	1058871,830	58826,213	1,225	0,230

Tableau 73 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur le F3 des voyelles du français

La Figure 93 de la page 207 illustre la variation du troisième formant des dix voyelles orales du français en fonction de la position dans le logatome qui est soit initiale, soit médiane, soit finale. La variation du F3 des voyelles postérieures n'est donnée qu'à titre indicatif et elle n'est pas prise en considération, comme nous l'avons expliqué dans la partie méthodologique. Nous y remarquons que le troisième formant varie peu selon la position dans le logatome.

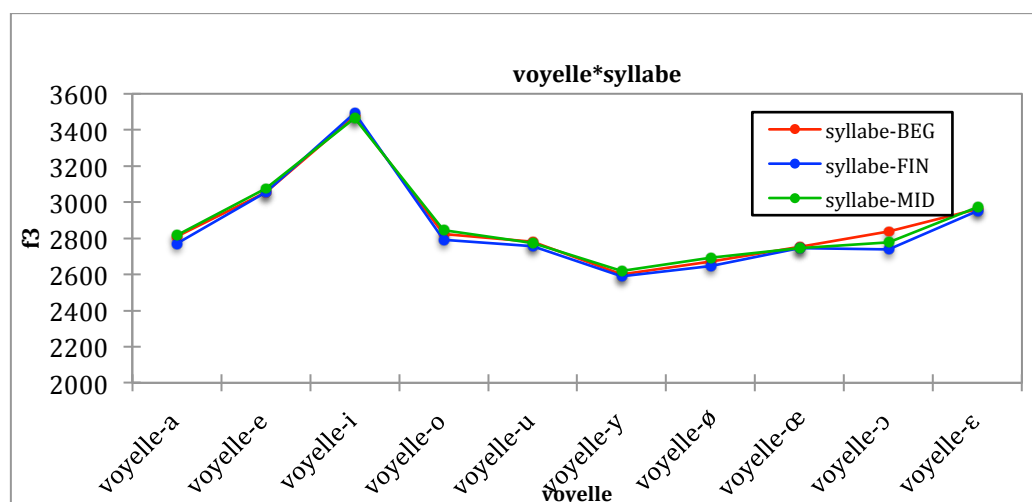


Figure 93 : F3 des voyelles orales du français en positions initiale (en rouge), médiane (en vert) et finale (en bleu), calculé à partir de la valeur prise à la moitié de la durée vocalique* 4 répétitions* 4 contextes confondus* 10 locutrices

Le résultat de l'analyse de variance ANOVA à deux facteurs ne confirme pas entièrement notre hypothèse de départ car les deuxième et troisième formants ne varient pas de manière significative selon si la voyelle avait été prononcée en positions initiale, médiane ou finale et le premier formant ne varie pas selon la position dans le logatome dans le cas de [y, u, o].

En revanche, il existe une petite différence (significative) entre les valeurs selon la position dans le logatome au niveau du premier formant dans le cas de [i, e, ε, ø, œ, ɔ, a]. Ces voyelles rencontrent des valeurs de F1 significativement différentes en position médiane par rapport aux positions initiale et/ou finale.

Pour résumer, nos résultats montrent que la position prosodique induisant un allongement final (la durée est en moyenne de 92, 93 et 133 ms pour respectivement les voyelles initiale, médiane et finale) n'affecte pas dans une grande mesure les valeurs des formants des voyelles insérées dans des logatomes trisyllabiques. Il faudrait bien sûr vérifier ce phénomène dans un corpus basé sur de vrais mots dans une tâche de lecture ou de production spontanée.

10.2 Réduction due à la position prosodique en tchèque

En tchèque, une réduction due à l'abrégement d'une voyelle en position inaccentuée d'un mot lexical n'est pas attendue (Flemming, 2006). En effet, la durée est utilisée au niveau phonologique et de ce fait elle ne peut pas être fortement corrélée à la réalisation de l'accent lexical initial, qui se manifeste essentiellement par des variations complexes de la fréquence fondamentale (Palková, 1997).

La littérature sur la réduction phonétique due à la position prosodique dans cette langue est restreinte. L'étude préliminaire de Dohalska *et al.* (2000) effectuée à partir de la parole spontanée et de la lecture de phrases par un locuteur tchèque n'a pas montré d'influence de la position prosodique sur la valeur des formants vocaliques. Notons cependant que Dubeda and Januska (2006) ont montré, à partir de la parole spontanée d'un locuteur tchèque pragoise, que les voyelles en position inaccentuée sont sujettes à de diverses déformations, telles que l'abrégement de voyelles longues, l'allongement de voyelles brèves, la contraction, l'élision, etc., plus souvent (de 9 %) que celles en position accentuée.

10.2.1 Variation de la durée suivant la position prosodique en tchèque

De même que pour le français, nous avons d'abord étudié l'effet de la position dans le logatome sur la durée vocalique. Le résultat est le même que pour le français et montre un allongement important de la dernière voyelle par rapport aux voyelles initiale et médiane. Ainsi, la durée vocalique moyenne des voyelles longues est de 212 ms (ET de 52 ms) en position finale et de 149 et 153 ms en positions initiale et médiane respectivement (avec les ET de 48 ms). Le rapport de durée est alors de 1 : 1,4 pour les voyelles longues en positions initiale-finale, de même que médiane-finale. La durée des voyelles brèves est en moyenne de 124 ms en position finale (ET de 33 ms) et de 82 et 86 ms en positions initiale et médiane respectivement (avec les ET de 28 ms). Ainsi, le rapport de durée est de 1 : 1,5 pour les voyelles en positions initiale-finale et de 1 : 1,4 en positions médiane-finale.

Afin de vérifier si les différences de durée sont significatives, nous avons conduit une analyse ANOVA à deux facteurs, voyelle et syllabe. Le facteur « Voyelle » comprend alors dix modalités, [ɪ, i:, ε, ε:, o, o:, u, u:, a, a:], et le facteur « Syllabe » a toujours trois modalités, initiale, médiane et finale.

La valeur ANOVA qui indique l'effet de la position dans le mot sur la durée de chacune des dix voyelles du tchèque est la suivante : $F(29, 9469) = 413$ avec $p < 0,05$. Le résultat se trouve dans le Tableau 74.

Durée	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	29	19187669,693	661643,783	412,605	< 0,0001
Erreur	9469	15184267,225	1603,577		

Tableau 74 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur la durée des voyelles tchèques

L'effet du facteur « Voyelle » est donné dans le par la valeur ANOVA $F(9, 9469) = 937$ avec $p < 0,05$, celui du facteur « Syllabe » est exprimé par la valeur $F(2, 9469) = 1669$ et $p < 0,05$ et enfin l'effet général des deux facteurs combinés « Voyelle*Syllabe » est celui donné par $F(18, 9469) = 11$ et $p < 0,05$. L'effet des facteurs étudiés est ainsi significatif.

Durée	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle	9	13522204,030	1502467,114	936,947	< 0,0001
syllabe	2	5351473,309	2675736,654	1668,605	< 0,0001
voyelle*syllabe	18	313992,354	17444,020	10,878	< 0,0001

Tableau 75 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur la durée des voyelles du tchèque

Les résultats complets du test Fisher, utilisé pour indiquer les groupes responsables de l'effet global de l'interaction « Voyelle*Syllabe » sur la durée des voyelles du tchèque, se trouvent dans les annexes à la page 69.

La Figure 94 montre l'évolution de la durée vocalique pour chaque voyelle séparément en fonction de sa position dans le mot (1 = initiale, 2 = médiane, 3 = finale).

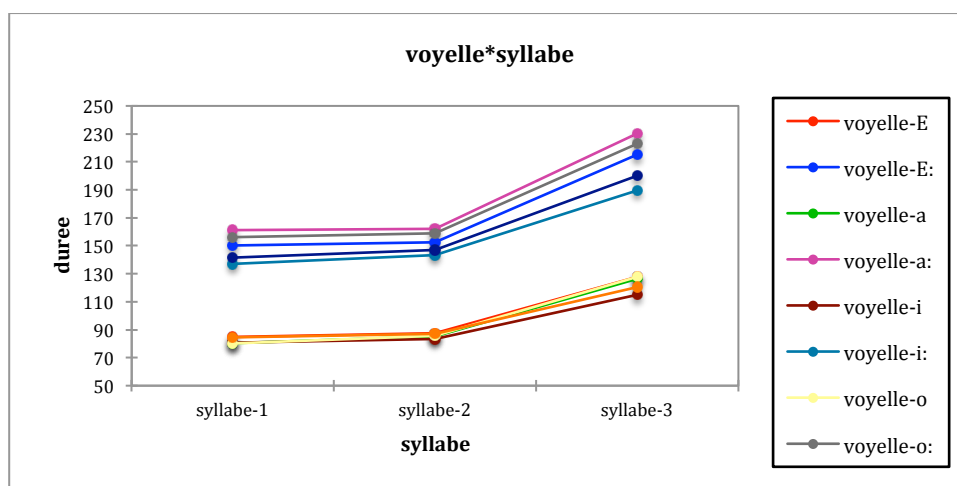


Figure 94 : Variation de durée des dix voyelles orales du tchèque [ɪ, i:, ɛ, ɛ:, o, o:, u, u:, a, a:] selon la position prosodique. Syllabe 1 = initiale, syllabe 2 = médiane, syllabe 3 = finale

Nous remarquons que les voyelles longues et brèves forment deux groupes bien séparés mais que les voyelles au sein de chaque groupe suivent la même tendance : les voyelles en syllabe initiale et médiane ont une durée comparable alors que les voyelles de la syllabe finale sont allongées de manière importante. La Figure 94 illustre ce comportement général, voyelles brèves et longues confondues.

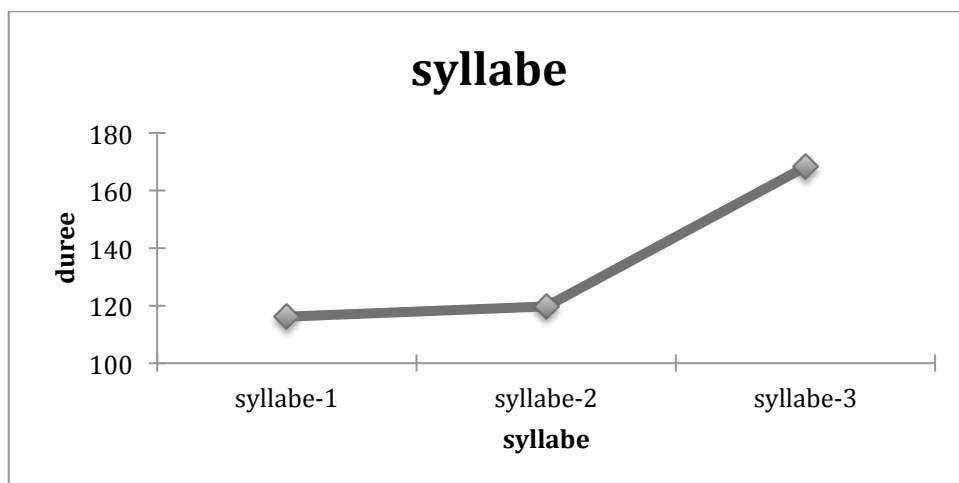


Figure 95 : Variation de durée vocalique en tchèque selon la position prosodique. Syllabe 1 = initiale, syllabe 2 = médiane, syllabe 3 = finale

Notre hypothèse, à savoir qu'il existe en tchèque une différence de durée vocalique selon la position de la voyelle dans le logatome, est confirmée car la voyelle finale est significativement plus longue que la voyelle initiale ou médiane. L'hypothèse nulle est rejetée.

10.2.2 Variation des formants suivant la position prosodique en tchèque

Nous avons ensuite observé l'effet de la position dans le logatome sur les formants vocaliques des voyelles du tchèque. Nous avons conduit une analyse ANOVA à deux facteurs, Voyelle et Syllabe. Le facteur « Voyelle » comprend dix modalités : [ɪ, i:, ɛ, ɛ:, o, o:, u, u:, a, a:] et le facteur « Syllabe » trois modalités : initiale, médiane et finale.

Les valeurs formantiques F1, F2, (F3) moyennes sont indiquées pour chaque voyelle et chaque position du logatome séparément dans le Tableau 76. Nous y remarquons que la valeur du premier formant de la syllabe finale est plus élevée que celui des syllabes initiale et médiane notamment dans le cas des voyelles brèves de grande ou moyenne aperture [ɛ, o, a] (cases grisées du tableau).

Formants Voyelles/ Syllabes			F1			F2			F3		
			Beg	Mid	Fin	Beg	Mid	Fin	Beg	Mid	Fin
Voyelles longues	i:	fermées	267 (29)	272 (38)	270 (33)	2691 (198)	2715 (201)	2717 (215)	3515 (219)	3517 (193)	3540 (193)
	u:		294 (41)	300 (41)	297 (45)	800 (129)	777 (144)	762 (111)			
	ɛ:	ouvertes/ moyennes	675 (81)	659 (82)	676 (66)	2056 (190)	2084 (196)	2062 (175)	2947 (170)	2939 (168)	2923 (154)
	o:		523 (82)	505 (79)	529 (83)	997 (111)	984 (111)	996 (103)			
	a:		894 (85)	881 (87)	893 (92)	1560 (148)	1539 (146)	1523 (146)	2751 (167)	2734 (186)	2735 (190)
Voyelles brèves	ɪ	fermées	386 (56)	380 (44)	412 (47)	2408 (215)	2417 (217)	2427 (211)	3018 (174)	3007 (162)	3011 (153)
	ʊ		327 (46)	337 (42)	358 (47)	839 (168)	850 (167)	833 (125)			
	ɛ	ouvertes/ moyennes	591 (105)	584 (104)	645 (75)	1999 (236)	2019 (234)	2019 (177)	2889 (179)	2887 (178)	2867 (158)
	o		460 (76)	454 (73)	524 (77)	1005 (126)	987 (127)	1012 (103)			
	a		770 (92)	755 (86)	801 (76)	1431 (178)	1396 (167)	1400 (145)	2760 (229)	2757 (235)	2712 (223)

Tableau 76 : Valeurs moyennes de F1 et F2 (toutes les voyelles), F3 (voyelles antérieures) des voyelles du tchèque, prononcées en syllabes initiale, médiane et finale dans des logatomes CVCVCVC où C = [p, t, k, f] (valeur prise à la moitié de la durée vocalique* 4 répétitions* 4 contextes confondus* 20 locutrices). Ecart type est entre parenthèses. Les variations selon la syllabe les plus amples sont grisées

	F1	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle		29	380547224,528	13122318,087	2645,060	< 0,0001
Erreur		9569	47472447,987	4961,067		

Tableau 77 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur le F1 des voyelles tchèques

La valeur ANOVA du facteur « Voyelle » est $F(9, 9569) = 8466$ avec $p < 0,05$. Celle du facteur « Syllabe » est $F(2, 9569) = 137$ avec $p < 0,05$ et finalement l'effet combiné des deux facteurs « Voyelle*Syllabe » est exprimé par $F(18, 9569) = 14$, $p < 0,05$ (Tableau 78). Les trois valeurs ANOVA élevées indiquent un effet significatif des facteurs étudiés.

F1	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle	9	377981769,177	41997974,353	8465,513	< 0,0001
syllabe	2	1359550,812	679775,406	137,022	< 0,0001
voyelle*syllabe	18	1205904,539	66994,697	13,504	< 0,0001

Tableau 78 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur le F1 des voyelles du tchèque

Le test Fisher précise à quels groupes est dû l'effet combiné des facteurs « Voyelle*Syllabe » sur le F1 et les résultats complets se trouvent dans la page 70 des annexes. Il indique que les changements de valeurs du F1 selon la position dépendent de la durée vocalique : alors qu'au niveau des voyelles brèves la variation est systématique et apparaît pour les cinq voyelles [a, ε, ɪ, o, u], (le premier formant de la voyelle finale est alors significativement plus élevé que celui des voyelles initiale et médiane dont le F1 est plus bas), au niveau des voyelles longues elle ne concerne que les voyelles de grande ou moyenne aperture [a:, ε:, o:] (le premier formant de la voyelle médiane est alors significativement plus bas que celui des voyelles initiale et finale). De cette façon, le premier formant des voyelles fermées longues [i:, u:] ne varie pas en fonction de la position dans le logatome. La Figure 96 montre l'évolution du F1 des dix voyelles monophthongues du tchèque, selon les positions initiale, médiane et finale.

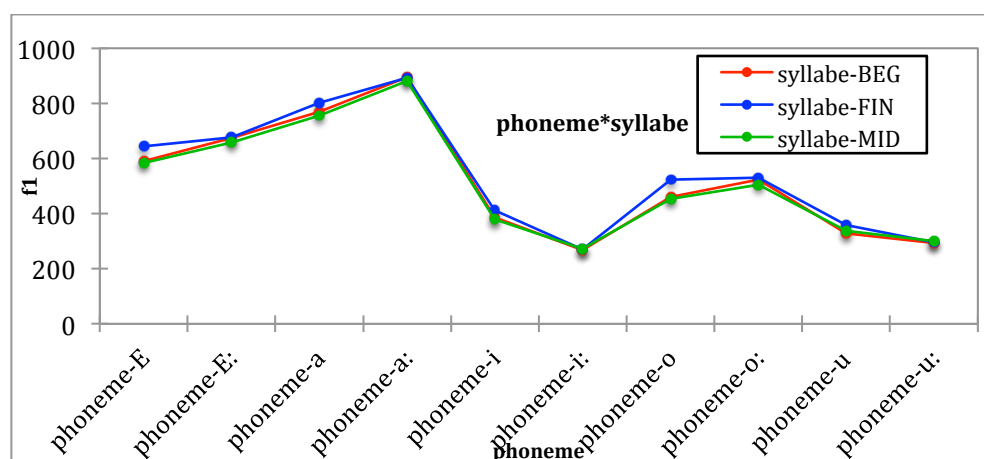


Figure 96 : F1 des voyelles monophthongues du tchèque en positions initiale (en rouge), médiane (en vert) et finale (en bleu), calculé à partir de la valeur centrale*4 répétitions*4 contextes confondus*20 locutrices

Le résultat est en concordance avec la littérature qui prévoit en position faible une baisse de la valeur du premier formant des voyelles ayant un F1 naturellement élevé, donc des voyelles ouvertes et mi-ouvertes (Wioland, 2005), de même que des changements formantiques moindres dans le cas des voyelles longues comparées aux voyelles brèves (Lindblom, 1963).

Au niveau du deuxième formant, la valeur ANOVA est $F(9569, 29) = 4989$ avec $p < 0,05$ (voir Tableau 79) et l'effet global de la position est ainsi significatif.

F2	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	29	4123598231,497	142193042,465	4989,283	< 0,0001
Erreur	9569	272713594,078	28499,696		

Tableau 79 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur le F2 des voyelles tchèques

La variable dépendante « Voyelle » a un grand effet sur la valeur du F2, qui est exprimé par $F(9, 9569) = 16072$ et $p < 0,05$. En revanche, la valeur ANOVA du deuxième facteur « Syllabe » est $F(2, 9569) = 0$ avec $p = 0,688$ (non significatif) et celle de l'interaction « Voyelle*Syllabe » est $F(18, 9569) = 2$ avec $p < 0,05$ (Tableau 80). L'effet de la position sur les valeurs du deuxième formant est ainsi moins grand que l'effet de la position sur le premier formant.

F2	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle	9	4122313610,492	458034845,610	16071,569	< 0,0001
syllabe	2	21322,896	10661,448	0,374	0,688
voyelle*syllabe	18	1263298,109	70183,228	2,463	0,001

Tableau 80 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur le F2 des voyelles du tchèque

Le test Fisher spécifie à quels groupes est dû l'effet global de l'interaction « Voyelle*Syllabe » sur le deuxième formant et les résultats complets sont indiqués à la page 70 des annexes. Il montre que le F2 de quatre voyelles brèves sur cinq [ɛ, ɪ, o, u] ne varie pas selon la position dans le logatome. Seul le F2 de la voyelle ouverte [a], comme nous le voyons à la Figure 97, est significativement plus élevé en position initiale qu'en positions médiane ou finale. En revanche, le deuxième formant des voyelles longues varie en fonction de la position dans le mot pour [a:] et [u:] où le F2 est plus élevé en position initiale qu'en position finale, pour [i:] où le F2 est significativement plus bas en position initiale qu'en position finale et pour [ɛ:] où le F2 est significativement plus bas en position initiale qu'en position médiane. Enfin, le deuxième formant de la voyelle [o:] ne change pas de valeur en fonction de la position prosodique. La Figure 97 illustre les mouvements du F2 selon la position de la voyelle dans le logatome.

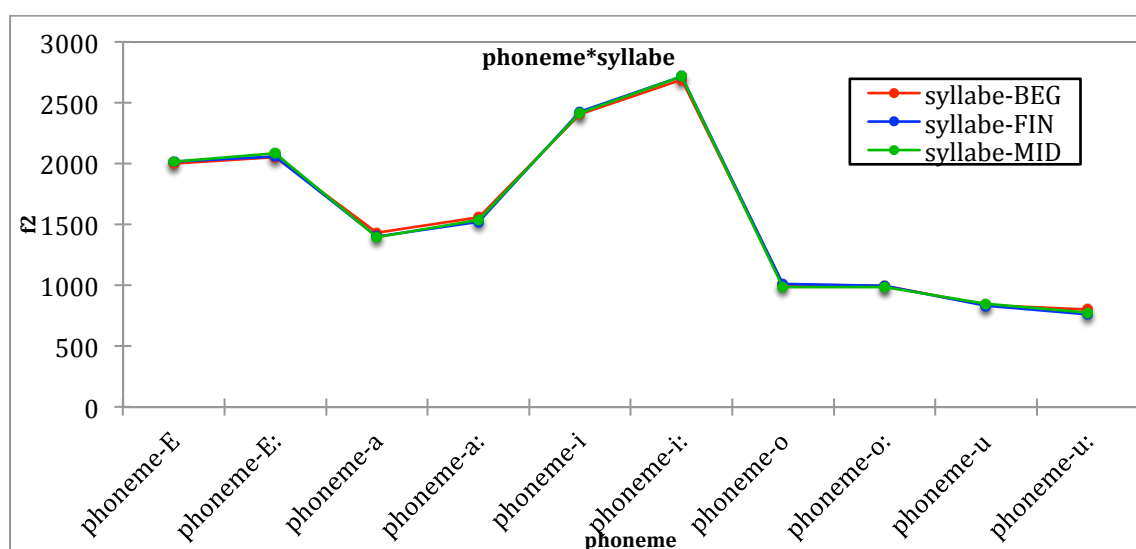


Figure 97 : F2 des voyelles monophthongues du tchèque en positions initiale (en rouge), médiane (en vert) et finale (en bleu), calculé à partir de la valeur prise à la moitié de la durée vocalique* 4 répétitions* 4 contextes confondus* 20 locutrices

Ainsi les variations du deuxième formant en fonction de la position dans le logatome sont aléatoires et concernent principalement les voyelles longues, ce qui est un résultat inattendu selon la littérature (Lindblom, 1963).

Enfin, le troisième formant est touché par l'effet de la position exprimé avec la valeur ANOVA $F(29, 9569) = 326$ avec $p < 0,05$ (Tableau 81).

F3	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
Modèle	29	482067006,929	16623000,239	325,730	< 0,0001
Erreur	9569	488335683,243	51033,095		

Tableau 81 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur le F3 des voyelles tchèques

L'effet du facteur « Voyelle » est donné par $F(9, 9569) = 1043$ avec $p < 0,05$, celui du facteur « Syllabe » est indiqué par $F(2, 9569) = 3$ avec $p < 0,05$ et enfin l'effet combiné des facteurs «Voyelle*Syllabe» sur le F3 est exprimé par la valeur ANOVA $F(18, 9569) = 3$ avec $p < 0,05$ (Tableau 82).

F3	DDL	Somme des carrés	Moyenne des carrés	F	Pr > F
voyelle	9	479032863,944	53225873,772	1042,968	< 0,0001
syllabe	2	328990,447	164495,224	3,223	0,040
voyelle*syllabe	18	2705152,537	150286,252	2,945	< 0,0001

Tableau 82 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur le F3 des voyelles du tchèque

Les résultats complets du test de Fisher indiquant à quels groupes est dû l'effet global de l'interaction « Voyelle*Syllabe » sur le F3 se trouvent à la page 71 des annexes. Le test a posteriori montre que le troisième formant de cinq voyelles [i, i:, e, e:, a:] sur six pour lesquelles on étudie les variations reste identique dans les trois positions du mot. La voyelle [a], comme illustré à la Figure 98, présente une valeur de F3 significativement plus basse en final du mot qu'en positions initiale ou médiane.

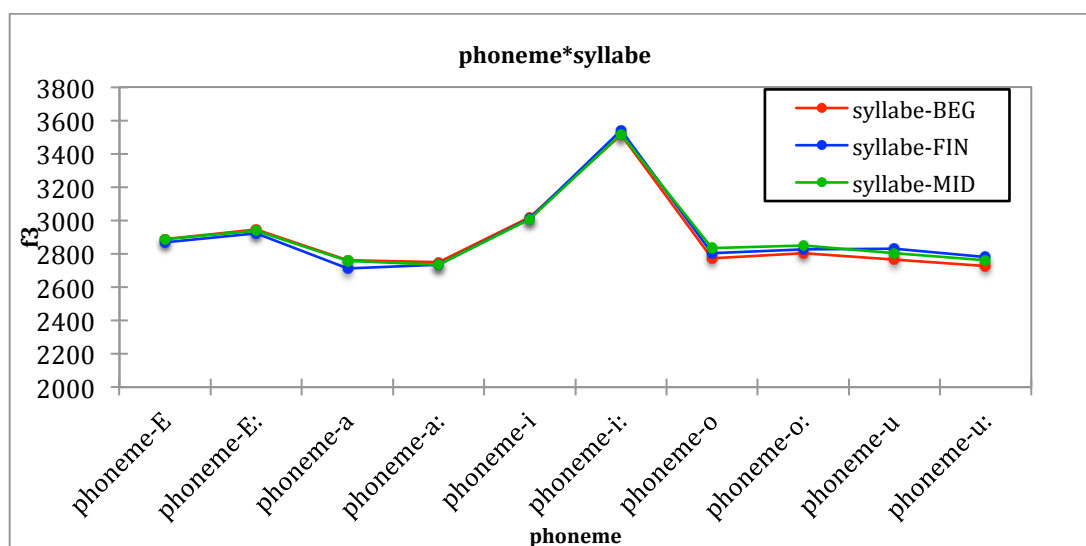


Figure 98 : F3 des voyelles monophtongues du tchèque en positions initiale (en rouge), médiane (en vert) et finale (en bleu), calculé à partir de la valeur prise à la moitié de la durée vocalique * 4 répétitions* 4 contextes confondus* 20 locutrices

Notre hypothèse à savoir que les valeurs formantiques varient selon la position de la voyelle dans le non-mot est confirmée et l'hypothèse nulle est rejetée. L'effet de la position dans le logatome sur les valeurs formantiques est plus grand pour le F1 que pour les formants supérieurs.

10.3 Réduction due à la position prosodique en français et en tchèque : Conclusion partielle

Nous relierons les différences duratives et spectrales des voyelles initiales et médianes par rapport aux voyelles finales des logatomes à l'allongement final.

Les variations de durée vocalique selon la position de la voyelle sont importantes en français et en tchèque : les voyelles longues finales du tchèque et les voyelles finales du français dans les logatomes trisyllabiques sont en moyenne 1,4 fois plus longues que les voyelles initiales et médianes. Les voyelles brèves du tchèque sont en moyenne 1,5 fois plus longues en position finale qu'en position initiale.

En revanche, les variations spectrales en fonction de la position prosodique sont petites, quoique l'analyse ANOVA montre leur significativité au niveau du premier formant pour les voyelles [i, e, ø, ε, œ, ɔ, a] du français et au niveau des trois formants pour quelques voyelles tchèques (avec le premier formant qui porte les variations les plus importantes).

Il est ainsi possible de dégager certaines régularités en ce qui concerne les patrons formantiques pour les voyelles produites en initiale du logatome, donc avec plus de force articulatoire et à la fin du logatome, donc avec un allongement prépausal par rapport aux voyelles en position médiane (plus brèves qu'en final et moins tendues qu'à l'initial).

Les résultats montrent un déplacement des voyelles dans l'espace acoustique F1/F2 essentiellement vertical, en tirant le plafond des triangles vocaliques vers le haut. En tchèque, le premier formant de toutes les voyelles brèves - mis à part le [ɪ] - baisse. Le [ɪ] s'avère une voyelle résistante à la réduction phonétique, ce qui est en accord avec la littérature. Le F1 des voyelles longues évolue en fonction de la position prosodique de façon négligeable. Il est à noter que la valeur maximale du premier formant (donc la valeur la moins réduite) est en tchèque toujours portée par la syllabe finale – allongée – et non pas par la syllabe initiale – accentuée. En français, les voyelles produites avec le F1 maximal sont soit celles de la syllabe finale soit celles de la syllabe initiale. Le F1 baisse systématiquement dans le cas des voyelles françaises mi-ouvertes [ɛ, œ, ɔ] et [a] sans pourtant menacer l'opposition entre les voyelles mi-ouvertes et mi-fermées. Le F1 des voyelles mi-fermées et fermées du français change minimalement suivant la position prosodique.

Nos résultats sont en accord avec un bon nombre d'études sur la réduction phonétique des voyelles due à leur abrégement (du fait de leur position faible ou du débit accéléré). L'étude classique de Lindblom (1963) sur des séquences CVC en suédois montre que la non-atteinte (ou bien dans les termes de l'auteur le *undershoot*) du F1 est plus grande pour la voyelle ouverte que pour les voyelles moyennes ou fermées. Ainsi plus la voyelle est produite rapidement, plus le F1 de [a] baisse. L'auteur reporte également la résistance de [i] à la réduction. Liljencrants and Lindblom (1972) fournissent une explication de la baisse du F1 des voyelles ouvertes abrégées : elle serait motivée par la minimisation de l'effort articulatoire qui est plus grand quand, dans la production de séquences CVC, le locuteur passe d'une totale occlusion vers une grande aperture, ce qu'exige l'articulation de [a]. Ce déplacement du plafond du triangle vocalique vers le haut pour les voyelles de courte durée par rapport aux voyelles longues se produit également dans les langues où la réduction phonétique est plus faible (elle ne résulte pas dans la substitution des voyelles par un schwa), comme en allemand (Mooshammer and Geng, 2008). Ce déplacement essentiellement vertical dans l'espace vocalique est ensuite reflété par la neutralisation des contrastes d'aperture qui est la plus courante dans les langues du monde. Flemming (2006, p. 5) précise à ce propos que : « *Vowel reduction primarily eliminates height contrasts, and only eliminates backness or rounding contrasts under restricted conditions.* » Comme exemple d'une langue avec suppression de contrastes d'aperture en syllabe inaccentuée, nous pouvons citer l'italien où la neutralisation concerne l'opposition entre les voyelles mi-fermées et mi-ouvertes ou

le russe dont l'inventaire vocalique se réduit de cinq voyelles [i, a, u, e, o] à trois, en supprimant les voyelles moyennes.

Même si la position médiane (par rapport aux positions initiale et/ ou finale) engendre une baisse du premier formant vocalique de [ɛ, œ, ɔ], notre résultat ne confirme pas la neutralisation de l'opposition des voyelles moyennes mi-ouvertes et mi-fermées en syllabe inaccentuée exposée par (Wioland, 2005).

La petite réduction spectrale des voyelles du français et du tchèque due à la position prosodique (non attendue pour le français par Delattre (1969) peut s'expliquer par :

- notre choix du corpus. Selon Durand (1985), les non-mots tendent à être hyperarticulés (et donc de longue durée) ce qui diminue la variation due à la position dans le logatome.
- un débit relativement lent. La durée vocalique moyenne dans toutes les positions du logatome (y compris en positions initiale et médiane) est supérieure à 90 ms pour le français et à 80 ms pour le tchèque
- le choix du corpus de Delattre (1969). En étudiant la réduction des voyelles du français, l'auteur explique les mouvements formantiques des voyelles non finales vers le centre acoustique par leur position inaccentuée. Cependant, en étudiant le corpus de Delattre plus en détail (voir le Tableau 83), nous expliquons les déplacements sur l'axe F1 et F2 plutôt par une coarticulation régressive avec la voyelle finale. Notons que Fagyal *et al.* (2002b) ont déjà démontré que la dilation se produit en français.

Nous remarquons dans le Tableau 83 que pour étudier la réduction des voyelles fermées du français, Delattre place ces dernières systématiquement devant des voyelles plus ouvertes et à l'inverse, les voyelles ouvertes dont le timbre est étudié sont placées devant les voyelles plus fermées. La seule exception concerne la paire « compose => composer » où [o] et [e] sont caractérisés par approximativement la même valeur du premier formant et où l'auteur observe effectivement le seul mouvement du deuxième formant de [o] qui augmente en se dirigeant vers le F2 de [e] (le F1 ne change pas). En ce sens, une plus petite réduction observée dans nos données peut s'expliquer par le contexte phonétique qui dans notre corpus reste strictement le même dans les trois positions du logatome.

Voyelle	Position accentuée	Position inaccentuée
i	fa'tigue	Fati'gant
e	Pre'ssé	Précé'dant
ɛ	De'teste	Détes'ter
a	Cons'tate	Consta'ter
ɔ	Pro'voque	Provo'quer
o	Com'pose	Compo'ser
u	Re'coupe	Recou'per
y	E'tude	Etu'dier
ø	Au'deux	Au deu'xième
œ	On'meuble	On meu'blait

Tableau 83 : Corpus de (Delattre, 1969) pour l'étude de la réduction des voyelles du français en position inaccentuée. La syllabe accentuée est précédée d'apostrophe. La voyelle dont le timbre est étudié est en italique. La voyelle finale imposant son lieu d'articulation est en gras

Nous sommes consciente qu'une étude plus approfondie avec un corpus basé sur la parole naturelle est nécessaire afin de confirmer nos résultats dégagés à partir de l'étude des logatomes.

11. Propriétés perceptives des voyelles : perception interlangue des voyelles du français par des Tchèques monolingues

Résumé : Le **chapitre onze** s'intéresse à l'assimilation perceptive des voyelles orales du français par dix Tchèques monolingues. Le résultat révèle que la similarité perceptive dépend de nouveau du contexte car en fonction de ce dernier, la même voyelle du français (par exemple le [y]) peut être soit perceptivement *nouvelle* (en contexte dental), soit perceptivement *similaire* à une voyelle du tchèque (au [ɪ] en isolation et en contexte palato-vélaire ainsi qu'au [i:] en contexte labial).

Afin d'établir la similarité perceptive entre les voyelles des deux langues en isolation et en contextes phonétiques [p, t, k], nous avons construit des tests d'identification interlangue. Dix auditeurs natifs tchèques de la région de Bohême ont identifié les dix voyelles orales du français [i, e, ε, y, ø, œ, u, o, ɔ, a] en termes de dix voyelles monophthongues du tchèque [ɪ, i:, ε, ε:, u, u:, o, o:, a, a:]. Ils ont attribué une note sur une échelle de 1 à 5 : 1 signifie que la voyelle est « très mal prononcée » alors que 5 correspond à « très bien prononcée ». La multiplication de la proportion de l'identification (pourcentage de l'identification divisé par 100) et de la note permet de calculer « l'index de correspondance » (le *fit index*), utilisé dans les travaux de Guion *et al.* (2000). Notons que dans l'interprétation de la similarité perceptive des voyelles du français et du tchèque, les paires des voyelles brèves-longues ε/ε:, o/o:, u/u:, a/a:, ayant la même qualité vocalique, comme nous l'avons démontré en 8.3.4, sont considérées ensemble.

La similarité perceptive des voyelles du français et du tchèque en fonction du *fit index* (FI) est établie comme suit:

Les voyelles du français “hautement similaires” aux voyelles du tchèques sont assimilées à une seule catégorie avec un *fit index* de 4 et plus.

Les voyelles du français *similaires* aux voyelles du tchèques sont assimilées à une ou plusieurs catégories dont une prédominante. Plus la valeur du *fit index* est basse, plus la similarité perceptive entre les deux voyelles est petite.

Les voyelles du français *nouvelles* sont assimilées à plusieurs catégories tchèques, dont aucune prédominante ou elles ne sont assimilées à aucune catégorie maternelle.

11.1 Voyelles isolées du français

Afin de vérifier la similarité perceptive des voyelles isolées du français [i, e, ε, y, ø, œ, u, o, ɔ, a] avec les voyelles du tchèque [ɪ, i:, ε, ε:, u, u:, o, o:, a, a:], nous avons construit un test d'identification interlangue. Les voyelles ont été enregistrées par trois locutrices françaises natives de Paris. Le test contient 120 stimuli dont 10 stimuli d'entraînement présentés en début de test. Le test est construit selon le protocole décrit dans la partie méthodologie du chapitre 6.

Le Tableau 84 récapitule les informations sur les locuteurs, auditeurs et stimuli du test de perception interlangue.

Type de test	Identification, interlangue
Stimuli	10 voyelles orales isolées du français [i, e, ε, y, ø, œ, u, o, ɔ, a], 12 séries (dont une série utilisée pour l'entraînement)
Auditeurs	10 Tchèques monolingues ; 1 Français natif
Locuteurs	3 Françaises parisiennes
Nb total de stimuli	120
Nb de stimuli d'entraînement	10
Nb de stimuli distracteurs	0
Echelle de justesse	1 – 5 1 très mal prononcée 2 plutôt mal prononcée 3 entre les deux 4 plutôt bien prononcée 5 très bien prononcée

Tableau 84 : Informations sur la construction et passation du test de perception interlangue des voyelles isolées

Le test commence par la consigne suivante (traduite en français):

Ceci est un test de perception.

Vous allez entendre des voyelles tchèques.

Indiquez la voyelle que vous avez entendue
en cliquant sur le bon mot.

Comment est-elle prononcée d'après vous ?

Sur une échelle de 1 à 5 choisissez la valeur :

5 signifie très bien prononcée

1 signifie très mal prononcée

3 signifie entre les deux

Une fois cliqué sur l'écran, le test est lancé. Les auditeurs sont invités à cliquer sur l'un des dix mots proposés, précédés par les graphèmes des voyelles. Les mots proposés pour l'identification des voyelles isolées /a, ε, ɪ, o, u, a:, ε:, i:, o:, u:/ sont respectivement had [ɦat] (le serpent), ven [ven] (dehors), byt [bit] (l'appartement), hop [ɦop] (hop), fuk [fuk] (égal), páv [pa:f] (le pan), péct [pɛ:tst] (cuire), pít [pi:t] (boire), pól [po:l] (le pôle) et púl [pu:l] (la moitié). Les auditeurs ont la possibilité de cliquer sur le bouton «Aucune des voyelles» (nous nous sommes inspirée ici du test de perception interlangue de Bohn et Flege (1990) pour signaler que le son perçu ne correspond à aucune voyelle tchèque. La Figure 99 à la page 219 montre l'interface graphique du test.

1 / 120

Vyberte správnou samohlásku,
kliknete na slovo, které obsahuje samohlásku,
co slyšíte.
Reknete, jak je samohláska vyslovena a
kliknete na OK.

ZADNA ZE SAMOHLÁSEK

a - had	e - ven	y - byt	o - hop	u - fuk
á - páv	é - péct	í - pít	ó - pól	ů - půl
1. Velmi špatně	2. Spis špatně	3. Neco mezi	4. Spis dobře	5. Vyborne vyslovena

Figure 99 : Interface graphique dans Praat du test de perception interlangue des voyelles isolées

Le test a été d'abord conduit par un locuteur français natif afin de vérifier l'identification des voyelles. Les consignes ont été traduites en français (en disant que l'auditeur allait entendre des voyelles françaises) et dans le choix de réponses, nous avons proposé les dix voyelles orales du français suivantes : « a – papa », « é – bébé », « i – pipi », « ô – beau », « O – bol », « E – belle », « u – pull », « ou – pou », « eu – peu », « oeu – poeur ». L'auditeur français natif a identifié toutes les voyelles isolées à 100 % avec la note de 5 (meilleure note).

Le résultat de l'identification des voyelles du français par dix Tchèques monolingues se trouve dans le Tableau 85 à la page 220. Nous y trouvons 1. les stimuli qui sont les dix voyelles orales du français (colonne 1), 2. les réponses des auditeurs (en termes de catégories tchèques) : les paires de voyelles brèves et longues sont considérées ensemble, mis à part le *ɪ/i* (colonne 2), 3. le pourcentage de l'identification de chaque voyelle française en termes de voyelle tchèque (colonne 3), 4. la note moyenne de l'échelle de confiance (colonne 4) et 5. le *fit index* (proportion de l'identification* note, colonne 5).

Voyelles FR	Voyelles TCH	%identification	Note	Fit index
a	a/ a:	100	4,4	4,4
ε	ε/ ε:	100	4	4
u	u/ u:	97,4	3,95	3,9
	AC	2,6		
i	i:	98,2	3,8	3,7
	ɪ	1,8	3,6	0,04
ɔ	o/ o:	89,1	3,6	3,2
	a/ a:	8,2	2,15	0,2
	AC	2,7		
o	u/ u:	73,4	4,35	3,2
	o/ o:	24,1	2,1	0,5
	AC	2,5		
e	ɪ	61	4	2,4
	ε/ ε:	33,2	2,2	0,7
	i:	2,5	2,6	0,1
	AC	3,3		
œ	ε/ ε:	85	2,55	2,3
	AC	15		
v	ɪ	71,5	2,3	1,6
	i:	21,1	2,9	0,6
	AC	7,4		
ø	ε/ ε:	38	1,55	0,6
	ɪ	13,2	1,5	0,2
	u/ u:	8,2	1,75	0,1
	AC	40,6	3,7	

Tableau 85 : Identification interlangue des voyelles du français par des Tchèques monolingues. Les résultats sont classés selon le *fit index*, calculé selon (Guion *et al.*, 2000). AC = aucune catégorie

Le Tableau 85 met en évidence la similarité perceptive des voyelles du français avec les voyelles du tchèque. Il montre que :

- La voyelle isolée du français [a] est toujours identifiée en termes de /a, a:/ isolé du tchèque, avec une note moyenne de 4,4. Il s'agit de la voyelle la plus *similaire* du point de vue perceptif.
- La voyelle française [ε] est également toujours identifiée en termes d'une seule qualité vocalique tchèque /ε, ε:/ avec une note moyenne 4. Il s'agit donc également d'une voyelle perceptivement « hautement similaire ».
- Deux autres voyelles françaises sont très *similaires* aux voyelles tchèques : le [u] et le [i]. Ainsi, le [u] est identifié en termes de /u, u:/ tchèque dans 97 % des réponses (dans 3 % des réponses il n'est perceptivement relié à aucune catégorie tchèque) avec une note moyenne de

3,95 et le [i] français est relié au /i:/ long tchèque dans 98 % des réponses avec une note moyenne de 3,8 (dans 2 % des réponses, il est identifié en termes de /ɪ/ bref).

- La voyelle française [ɔ] est classée cinquième du point de vue de similarité perceptive. Elle est perceptivement reliée à une catégorie prédominante qui est le /o, o:/ tchèque, dans 89 % des réponses, avec une note moyenne de 3,6, mais également au /a, a:/ tchèque dans 8 % des réponses avec une note de 2,12 et à aucune catégorie maternelle dans 3 % des réponses.
- Le [o] fermé du français est identifié en termes de /u, u:/ tchèque dans 73 % des réponses avec une note élevée de 4,35 (la deuxième meilleure note). Cependant, il est également identifié comme /o, o:/ dans 24 % des réponses mais avec une note basse de 2,1. Dans 3 % des réponses, le [o] français n'est perceptivement relié à aucune catégorie tchèque.
- La voyelle classée septième selon le *fit index* exprimant la similarité perceptive est le [e] français. Il est perceptivement *similaire* au /ɪ/ bref tchèque dans 61 % des réponses avec une note forte de 4 mais également au /ɛ, ɛ:/ tchèque dans 33 % des réponses avec une note faible de 2,2. Dans 3 % des réponses, le [e] français est identifié en termes de /i:/ (avec une note de 2,6) et également en termes d'aucune catégorie tchèque.

Ils restent trois voyelles du français, retranscrites par des symboles phonétiques qui ne sont pas utilisés pour la transcription des voyelles tchèques, les [œ], [y] et [ø]. La similarité de ces voyelles isolées aux voyelles tchèques est la suivante :

- Le [œ] isolé du français est assimilé à une seule qualité vocalique tchèque, qui est le /ɛ, ɛ:/, mais avec une note basse de 2,55 sur 5 et dans 15 % des réponses la voyelle n'est perceptivement reliée à aucune catégorie tchèque. Il s'agit donc d'un son *similaire*.
- Le [y] est le plus souvent assimilé au /ɪ/ tchèque, dans 72 % des réponses et avec une note de 2,3. Dans 21 % des réponses, le [y] est relié au /i:/ long tchèque avec une note de 2,9 et dans 7 % des réponses, il n'est perceptivement relié à aucune voyelle tchèque. C'est une voyelle perceptivement *similaire*.
- Enfin, le [ø] français est perceptivement relié à plusieurs qualités vocaliques tchèques, les /ɛ, ɛ:/, le /ɪ/ et les /u, u:/, avec une note moyenne respectivement de 1,55, 1,5 et 1,75. Cependant dans 41 % des réponses, il n'est relié à aucune catégorie tchèque. La voyelle française [ø] est donc véritablement *nouvelle*.

Selon le test de perception interlangue, la similarité des voyelles isolées du français aux voyelles tchèques baissent progressivement de : [a] > [ɛ] > [u] > [i] > [ɔ] > [o] > [e] > [œ] > [y] > [ø].

11.2 Voyelles du français en contextes p, t, k

Afin d'établir la similarité perceptive des voyelles du français en contexte aux voyelles du tchèque, nous avons construit un deuxième test dont les stimuli sont des syllabes symétriques CVC où C correspond aux consonnes [p, t, k] et V aux dix voyelles orales du français [i, e, ɛ, y, ø, œ, u, o, ɔ, a]. Nous avons ainsi créé dix stimuli du type pVp, tVt et kVk, chacun répété quatre fois. Les 120 stimuli sont précédés de 10 stimuli d'entraînement et complétés par 5 stimuli distracteurs qui sont des mots monosyllabiques tchèques sans noyau vocalique : « smrk », « mrk », « krk », « blb », « srk ». Dans les boutons de réponse, seule la voyelle est indiquée, sans mot de référence : « a », « á », « e », « é », « i », « í », « o », « ó », « u », « ú ». Notons que les graphèmes tchèques sont identiques aux symboles phonétiques traditionnellement utilisés pour leur transcription. Les stimuli ont été enregistrés par une Française parisienne et l'échelle de justesse s'étend de 1 à 5 où 5 correspond à la meilleure note.

Le Tableau 86 récapitule les informations sur le test interlangue d'identification de voyelles en contexte.

Type de test	Identification, interlangue
Stimuli	10 voyelles orales du français [i, e, ε, y, ø, œ, u, o, ɔ, a] en contextes symétrique [p, t, k], 4 séries pour chaque contexte
Auditeurs	10 Tchèques monolingues ; 1 Français natif
Locuteurs	1 Française parisienne
Nb total de stimuli	135
Nb de stimuli d'entraînement	10
Nb de stimuli distracteurs	5
Echelle de justesse	1 – 5 1 très mal prononcée 2 plutôt mal prononcée 3 entre les deux 4 plutôt bien prononcée 5 très bien prononcée

Tableau 86 : Informations sur la construction et passation du test de perception interlangue des voyelles françaises en contextes pVp, tVt, kVk

Afin de vérifier l'identification des voyelles orales du français en contextes symétriques pVp, tVt, kVk par un Français natif, nous avons adapté le test en le traduisant en français. Le résultat détaillé se trouve en annexe du chapitre 11 (pages 85 et 86 des annexes).

En comparant ces résultats avec l'identification des voyelles isolées chez l'auditeur français natif, nous constatons que :

- Les voyelles isolées du français sont plus facilement identifiables que les voyelles en contexte.
- Les voyelles en contexte identifiées avec une note de qualité inférieure à 5 sont 1. les voyelles moyennes [e, œ] dans tous les contextes 2. [ε] en contextes [t, k], 3.[ɔ] en contexte [t]. Notons que l'attribution d'une note de qualité plus basse aux [ɔ, œ, ε] qu'à d'autres voyelles ne s'explique pas par la loi de position qui autorise l'insertion des voyelles mi-ouvertes en syllabe fermée.
- La voyelle la moins facilement identifiable dans tous les contextes est le [e] français, ce qui peut s'expliquer par les règles orthoépiques. Son identification est maximale en contexte labial (où elle est une fois confondue avec le /i/), puis en contexte dental (où son identification en termes de /e/ visé est accompagnée d'une petite note de qualité).

Ainsi en prenant en compte une relative difficulté à identifier les voyelles moyennes /ɔ, œ, ε, e/ par un natif, nous avons proposé le même test à dix auditeurs tchèques monolingues qui ont identifié les voyelles du français en termes de catégories tchèques. Les résultats du test d'identification des dix voyelles orales insérées en contexte symétrique labial se trouvent dans le Tableau 87 à la page 224. Ce tableau montre la similarité perceptive des voyelles du français aux voyelles du tchèque en contexte labial nous permet d'avancer les constats suivants :

- Quatre voyelles du français sont "hautement similaires" aux voyelles du tchèque. Le [u] français est "hautement similaire" aux [u, u:] tchèques (dans 100 % des réponses avec une note

moyenne de 4,6), le [o] aux [u/u:] tchèque (dans 97,5 % des réponses avec une note de 4,7, un stimuli est identifié comme /o:/), le [i] au [i:] tchèque (dans 100 % des réponses avec une note de 4,5) et enfin le [ɛ] aux [ɛ, ɛ:] (dans 90 % des réponses avec une note de 4,6, un stimuli est identifié en termes de /a:/ et trois autres comme /ɪ/).

- La voyelle [e] du français est *similaire* au [ɪ] tchèque dans 85 % des réponses avec une note moyenne de 4,5 mais elle est également reliée au /i:/ dans 10 % des réponses et au /ɛ/ dans 5 % des réponses.
- La voyelle française [a] est *similaire* aux [a, a:] dans 85 % des réponses avec une note de 4,4 et elle est aussi identifiée en termes de /ɛ/ɛ:/ dans 15 % des réponses.
- Le [œ] entretient une similarité perceptive avec le [ɛ/ɛ:] (cette assimilation se produit selon 80% des réponses avec une note de 3,8 quoique [œ] soit également relié une fois au /ɪ/, une fois au /o/ et dans 15 % des réponses il ne correspond à aucune catégorie tchèque).
- Le [ɔ] est interprété par la catégorie prédominante /o/o:/ dans 70 % des réponses avec une note de 3,6. Les auditeurs l'identifient également en termes de /a/ dans 15 % des réponses. Il est ensuite deux fois identifié en tant que /ɛ/, une fois en tant que /u/ et trois réponses indiquent que [ɔ] ne correspond à aucune catégorie tchèque.
- Le [y] est interprété soit en termes de /i:/ (dans 65 % des réponses avec une note de 3,3), soit en termes de /ɪ/ (dans 30 % des réponses avec une note de 3,3) et deux réponses indiquent que [y] ne correspond à aucune catégorie tchèque. Il s'agit d'un son peu *similaire*.
- Enfin, la voyelle [ø] est très peu *similaire* car elle est perceptivement reliée à de nombreuses catégories tchèques (dont une prédominante) : aux /ɛ/ɛ:/ dans 45 % des réponses avec une note de 2,8, au /ɪ/ dans 20 % des réponses avec une note de 3,4, aux /u/u:/ dans 22,5 % des réponses avec une note de 3, un stimulus est identifié comme /o:/ et quatre réponses (10 %) indiquent que [ø] ne correspond à aucune catégorie tchèque.

Ainsi, le classement des voyelles françaises selon leur similarité perceptive avec les voyelles tchèques en contexte labial, du plus au moins, est le suivant : [u] > [o] > [i] > [ɛ] > [e] > [a] > [œ] > [ɔ] > [y] > [ø]. En contexte labial, aucune voyelle du français n'est perceptivement véritablement *nouvelle* car malgré une faible similarité perceptive, le [ø] est assimilé à une catégorie prédominante qui est le /ɛ/ɛ:/ tchèque.

Voyelles FR	Voyelles TCH	%identification	Note	Fit index
u	u/ u:	100	4,6	4,6
o	u/u:	97,5	4,7	4,6
	o:	2,5	1	0,1
i	i:	100	4,5	4,5
ε	ε/ε:	90	4,6	4,1
	ɪ	7,5	4	0,3
	a:	2,5	4	0,1
e	ɪ	85	4,5	3,8
	ε	5	4	0,2
	i:	10	4,7	0,5
a	a/a:	85	4,4	3,7
	ε/ε:	15	4,8	0,7
œ	ε/ ε:	80	3,8	3
	ɪ	2,5	2	0,1
	o	2,5	3	0,1
	AC	15		
ɔ	o/o:	70	3,6	2,5
	a	15	4,3	0,6
	ε	5	3,5	0,2
	u	2,5	2	0,1
	AC	7,5		
y	i:	65	3,3	2,1
	ɪ	30	3,3	1
	AC	5		
ø	ε/ ε:	45	2,8	1,3
	ɪ	20	3,4	0,7
	u/u:	22,5	3	0,7
	o:	2,5	4	0,1
	AC	10		

Tableau 87 : Identification interlangue des voyelles du français en contexte labial (4 répétitions) par 10 Tchèques monolingues. Les résultats sont classés selon le *fit index*, calculé selon (Guion *et al.*, 2000). AC = aucune catégorie

Les résultats de l'identification interlangue des voyelles en contexte symétrique dental se trouvent ensuite dans le Tableau 88 à la page 226. Il montre que :

- Trois voyelles du français sont "hautement similaires" : le [u], le [ɛ] et le [o]. Le [u] est identifié dans 100 % des réponses comme /u/u:/ avec une note moyenne de 4,4. Le [ɛ] est identifié dans 95 % des réponses comme /ɛ/ɛ:/ avec une note de 4,5 et 2 stimuli sont identifiés comme /ɪ/. Le [o] est identifié dans 97,5 % des réponses comme /u/u:/ tchèque avec une note moyenne de 4,4 et un stimulus est catégorisé /o/.
- La voyelle [e] quant à elle, est très *similaire* au [ɪ] tchèque (dans 90 % des réponses avec une note de 4,3), mais elle est aussi perceptivement reliée au /ɛ/ dans 7,5 % des réponses et une réponse indique que [e] ne correspond à aucune catégorie tchèque.
- La voyelle [œ] est *similaire* aux [ɛ/ɛ:] tchèque dans 87,5 % des réponses avec une note de 3,9, deux réponses la relient avec le /ɪ/ tchèque et trois réponses indiquent qu'elle ne correspond à aucune catégorie tchèque.
- Le [ɔ] est perceptivement relié au /o/o:/ tchèque dans 77,5 % des réponses avec une note de 4,1 et au /ɛ/ dans 12,5 %. Une réponse le relie au /a/ et trois réponses indiquent qu'il ne correspond à aucune catégorie tchèque.
- Le [i] est *similaire* au [i:] tchèque dans 97,5 % des réponses mais avec une note faible de 3,1. Une réponse indique qu'il ne correspond à aucune catégorie tchèque.
- Le [a] en contexte dental est relié soit aux /a/a:/ tchèques, dans 60 % des réponses avec une note de 4,4, soit aux /ɛ/ɛ:/, dans 40 % des réponses et avec une note de 3,9.
- Le [ø] est assimilé aux /u/u:/ tchèques dans 80 % des réponses et avec une note de 3, au /ɪ/ dans 15 % des réponses, une réponse le relie avec le /o/ tchèque et une réponse indique qu'il ne correspond à aucune voyelle du tchèque.
- Enfin, la voyelle [y] est perceptivement *nouvelle* (similarité 0). Elle n'est assimilée à aucune catégorie prédominante. Elle est perceptivement reliée soit au /i:/, dans 32,5 % des réponses avec une note de 3,6, soit au /u/u:/ dans 37,5 % des réponses avec une note de 2,3 ou encore au /ɪ/ dans 25 % des réponses. Deux réponses indiquent qu'elle ne correspond à aucune voyelle du tchèque.

Le classement des voyelles du français selon leur degré de similarité en contexte dental avec les voyelles tchèques est le suivant (du + au -) : [u] > [ɛ] > [o] > [e] > [œ] > [ɔ] > [i] > [a] > [ø] > [y]. Seule la voyelle [y] est perceptivement nouvelle.

Voyelles FR	Voyelles TCH	%identification	Note	<i>Fit index</i>
u	u/u:	100	4,4	4,4
ε	ε/ε:	95	4,5	4,4
	ɪ	5	3	0,2
o	u/u:	97,5	4,4	4,3
	o	2,5	4	0,1
e	ɪ	90	4,3	3,9
	ε	7,5	4,3	0,3
	AC	2,5		
œ	ε/ε:	87,5	3,9	3,4
	ɪ	5	2	0,1
	AC	7,5		
ɔ	o/o:	77,5	4,1	3,2
	ε	12,5	2,6	0,3
	a	2,5	3	0,1
	AC	7,5		
i	i:	97,5	3,1	3
	AC	2,5		
a	ε/ε:	60	4,4	2,6
	a/a:	40	3,9	1,6
ø	u/u:	80	3	2,4
	ɪ	15	2,7	0,4
	o	2,5	2	0,1
	AC	2,5		
y	i:	32,5	3,6	1,2
	u/u:	37,5	2,3	0,9
	ɪ	25	2,2	0,5
	AC	5		

Tableau 88 : Identification interlangue des voyelles du français en contexte dental (4 répétitions) par 10 Tchèques monolingues. Les résultats sont classés selon le *fit index*, calculé selon (Guion *et al.*, 2000). AC = aucune catégorie

Enfin, la similarité entre les voyelles du français et du tchèque en contexte palato-vélaire est démontrée dans le Tableau 89 de la page 228. Il indique que :

11. Propriétés perceptives des voyelles : perception interlangue des voyelles du français par des Tchèques monolingues

- Les voyelles “hautement similaires” sont le [o, u, ɔ, ɛ].
- Le [o] français est identifié en tant que /u/u:/ tchèques dans 97,5 % des réponses avec une note de 4,7. Une réponse le relie au /o/ tchèque.
- Le [u] français est identifié dans 100 % des réponses en termes de /u/u:/ tchèques, avec une note moyenne de 4,5. Il est intéressant de noter qu’en syllabe kV_k, le [o] est perceptivement plus proche du [u] tchèque que le [u] français.
- Le [ɔ] est relié aux /o/o:/ dans 95 % des réponses avec une note de 4,2, une réponse le relie au /u/ et une autre à aucune catégorie tchèque.
- Le [ɛ] est interprété en tant que /ε/ε:/ dans 97,5 % des réponses avec une note de 4,1 et une réponse le relie au /i/ tchèque.
- Ensuite, la voyelle [œ] est *similaire* au [ε/ε:] dans 92,5 % des réponses avec une note de 3,8 et trois réponses indiquent qu’elle ne correspond à aucune voyelle du tchèque.
- Le [e] est identifié en termes de /ɪ/ dans 77,5 % des réponses avec une note de 4,3 et en termes de /ε/ε:/ dans 22,5 % des réponses avec une note de 4.
- Le [i] est identifié dans 75 % des réponses en tant que /i:/ long, avec une note de 4,2 et dans 25 % des réponses en tant que /ɪ/ bref avec la même note de 4,2.
- La voyelle [a] en contexte palato-vélaire est reliée aux /ε/ε:/ tchèque dans 67,5 % des réponses avec une note de 3,7 et aux /a/a:/ dans 32,5 % des réponses avec une note de 2,9.
- La voyelle [y] est peu *similaire* : les auditeurs tchèques la relient au /ɪ/ tchèque dans 52,5 % des réponses avec une note de 2,5, au /i:/ dans 25 % des réponses avec une note de 3,7, mais aussi au /u/u:/ dans 17,5 % des réponses avec une note de deux. Deux réponses indiquent qu’elle ne correspond à aucune voyelle du tchèque.
- Enfin, la voyelle [ø] est *nouvelle* car elle n’est reliée à aucune catégorie tchèque prédominante. Les Tchèques l’interprètent en tant que /ε/ε:/ dans 35 % des réponses avec une note de 3,2, en tant que /u/u:/ dans 35 % des réponses avec une note de 2,8, en tant que /ɪ/ dans 22,5 % des réponses avec une note de 2,9. Une réponse la relie au /o/ est deux réponses indiquent qu’elle ne correspond à aucune voyelle du tchèque.

Ainsi, la similarité perceptive des voyelles françaises aux voyelles tchèques en contexte palato-vélaire décroît de : [o] > [u] > [ɔ] > [ɛ] > [œ] > [e] > [i] > [a] > [y] > [ø]. Seule la voyelle [ø] est véritablement *nouvelle* car elle n’est perceptivement reliée à aucune catégorie tchèque prédominante.

Voyelles FR	Voyelles TCH	%identification	Note	<i>Fit index</i>
o	u/u:	97,5	4,7	4,6
	o	2,5	2	0,1
u	u/u:	100	4,5	4,5
ɔ	o/o:	95	4,2	4
	u	2,5	2	0,1
	AC	2,5		
ɛ	ɛ/ɛ:	97,5	4,1	4
	ɪ	2,5	3	0,1
œ	ɛ/ɛ:	92,5	3,8	3,5
	AC	7,5		
e	ɪ	77,5	4,3	3,3
	ɛ/ɛ:	22,5	4	0,9
i	i:	75	4,2	3,2
	ɪ	25	4,2	1,1
a	ɛ/ɛ:	67,5	3,7	2,5
	a/a:	32,5	2,9	0,9
y	ɪ	52,5	2,5	1,3
	i:	25	3,7	0,9
	u/u:	17,5	2	0,4
	AC	5		
ø	ɛ/ɛ:	35	3,2	1,1
	u/u:	35	2,8	1
	ɪ	22,5	2,9	0,7
	o	2,5	1	0,1
	AC	5		

Tableau 89 : Identification interlangue des voyelles du français en contexte palato-vélaire (4 répétitions) par 10 Tchèques monolingues. Les résultats sont classés selon le *fit index*, calculé selon (Guion *et al.*, 2000). AC = aucune catégorie

En comparant l'identification par un Français natif et dix Tchèques monolingues des voyelles du français prononcées en isolation et en contextes phonétiques, nous constatons que :

- les voyelles [y, ø] sont facilement identifiables en termes de voyelles françaises et difficilement identifiables en termes de voyelles tchèques,
- les voyelles [i, u, o] sont facilement identifiables en termes de voyelles françaises et tchèques. Notons que le [o] français est identifié essentiellement en tant que /u/u:/ tchèques,
- la voyelle [a] est facilement identifiable en termes de voyelle française dans tous les contextes et en termes de voyelle tchèque uniquement en contexte labial. En contextes dental et palato-vélaire, le [a] français est assimilé à deux catégories tchèques : /ε/ε:/ (catégorie prédominante) et /a/a:/ ;
- la voyelle [e] dans les suites [pep], [tet] et [kek] est identifiée par les Tchèques avec un *fit index* respectivement de 3,8, puis 3,9 et 3,3 en tant que /ɪ/ tchèque et par le Français avec un *fit index* respectivement de 2,25, puis 2,75 et 3,75 en tant que /e/ français. Les Tchèques ont alors plus de facilité à catégoriser la voyelle de la suite [pep] et [tet] en termes de voyelle tchèque que le Français en termes de voyelle française. Ce phénomène s'explique par les règles orthoépiques selon lesquelles [e] ne doit pas apparaître en français standard en syllabe finale fermée.

La comparaison des propriétés perceptives des voyelles du français et du tchèque a mis en évidence que leur similarité phonétique varie fortement en fonction du contexte.

Conclusion partielle sur la similarité des voyelles du français non-méridional et du tchèque de Bohême

Nous avons établi la similarité entre les voyelles françaises et tchèques des points de vue « symbolique » (en comparant les symboles phonétiques utilisés par les auteurs), perceptif (en conduisant des tests de perception interlangue) et acoustique (en comparant les patrons formantiques). Le Tableau 90 résume les résultats des tests de perception et de l'analyse acoustique.

Voyelle/contexte		0		P		T		K	
FR	TCH	ET	FI	ET	FI	ET	FI	ET	FI
i	i:	- 2	3,7	- 1,5	4,5	- 1,5	3	- 1,5	3,2
e	ɪ	- 2	2,4	- 1	3,8	- 1	3,9	- 1,5	3,3
	ɛ/ɛ:	+ 2	0,7	+ 2	0,2	+ 2	0,3	- 1,5	0,9
ɛ	ɛ/ɛ:	- 2	4	- 1	4,1	- 1	4,4	- 1	4
a	a/a:	- 1	4,4	- 1	3,7	- 2	1,6	+ 2	0,9
	ɛ/ɛ:	+ 2	/	+ 2	0,7	- 2	2,6	- 2	2,5
u	u/u:	- 1,5	3,9	- 1,5	4,6	- 1,5	4,4	- 1,5	4,5
o	o/o:	+ 2	0,5	- 1	0,1	- 1,5	0,1	- 1	0,1
	u/u:	- 2	3,2	+ 2	4,6	+ 2	4,3	+ 2	4,6
ɔ	o/o:	- 1,5	3,2	+ 2	2,5	+ 2	3,2	+ 2	4
	a/a:	+ 2	0,2	- 2	0,6	+ 2	0,1	+ 2	/
y	ɪ	+ 2	1,6	+ 2	1	+ 2	0,5	+ 2	1,3
	i:	+ 2	0,6	+ 2	2,1	+ 2	1,2	+ 2	0,9
	u/u:	+ 2	/	+ 2	/	+ 2	0,9	+ 2	0,4
ø	ɛ/ɛ:	+ 2	0,6	+ 2	1,3	+ 2	/	+ 2	1,1
	u/u:	+ 2	0,1	+ 2	0,7	+ 2	2,4	+ 2	1
	ɪ	+ 2	0,2	+ 2	0,7	+ 2	0,4	+ 2	/
œ	ɛ/ɛ:	+ 2	2,3	- 2	3	- 1	3,4	- 2	3,5

Tableau 90 : Similarité acoustique (exprimée par le nombre d'écarts types ET des formants F1/F2, (F2/F3)) et perceptive (exprimée par le *fit index*) entre les voyelles du français (FR) et du tchèque (TCH) en contextes nul (0), labial (p), dental (t) et palato-vélaire (k). Voyelles *nouvelles* et voyelles "hautement similaires" grisées (gris foncé et gris clair respectivement)

Le Tableau 90 montre de quelle(s) voyelle(s) tchèque(s) se rapproche le plus chacune des dix voyelles orales du français, acoustiquement et perceptivement. La colonne « ET » indique la distance acoustique (en nombre d'écart types) entre les formants F1, F2, (F3 - des voyelles antérieures) des voyelles du français et du tchèque. La colonne « FI » (*Fit Index*, allant de 0 à 5) reflète ensuite la similarité perceptive globale (pourcentage d'identification et note moyenne confondus).

En interprétant les résultats dans leur ensemble, les voyelles *nouvelles* sont celles étant retranscrites avec des symboles phonétiques non exploités en tchèque, ayant une similarité perceptive zéro ou étant assimilées perceptivement à plusieurs catégories tchèques dont aucune prédominante et se trouvant dans des zones acoustiquement vides, c'est-à-dire à plus de deux écarts types de n'importe quelle voyelle tchèque. Le Tableau 90 nous permet d'avancer les constats suivants :

- Seules deux voyelles peuvent être considérées comme *nouvelles* : le [ø] prononcé isolément et en contexte symétrique palato-vélaire, puis le [y] en contexte symétrique dental (cases grises foncé). Même si ces voyelles se trouvent dans les autres contextes dans des zones F1/F2 acoustiquement vides, elles sont perceptivement reliées à une catégorie tchèque prédominante (assimilées avec un *fit index* néanmoins faible) ce qui nous empêche de dire qu'elles sont globalement complètement *nouvelles*.
- Deux voyelles sont "hautement similaires" (selon les trois critères utilisés) aux voyelles tchèques (cases grises clair) : le [ɛ] dans les suites [pɛp], [tɛt] et [kɛk] (au /ɛ/ɛ:/ tchèques) et le [a] en isolation. Elles sont retranscrites avec un symbole phonétique utilisé en tchèque, leurs formants F1, F2 et F3 se réalisent à moins d'un écart type de ceux de leur équivalent tchèque et enfin elles sont identifiées avec un *fit index* de quatre et plus. Il est à noter que le [a] en contexte labial est acoustiquement « hautement similaire » au [a/a:] tchèque mais il s'en éloigne davantage perceptivement et que le [ɛ] en isolation est perceptivement « hautement similaire » au [ɛ/ɛ:] tchèque mais il s'en éloigne davantage au niveau du patron formantique.
- Deux autres voyelles sont "hautement similaires" du point de vue perceptif et très *similaires* du point de vue acoustique: il s'agit du [i] dans la suite [pip] et du [u] prononcé dans les suites [pup], [tut] et [kuk].
- Ensuite, le [e] est acoustiquement « hautement similaire » et perceptivement *similaire* au [ɪ] tchèque dans les suites [pɛp] et [tɛt]. Notons alors que les symboles utilisés pour retranscrire les deux sons proches ([e] en français versus [ɪ] en tchèque) sur le plan acoustico-perceptif divergent. De même, la voyelle [o] dans les suites [pop], [tot] et [kok] est « hautement similaire » perceptivement, mais pas acoustiquement, à la voyelle tchèque [u]. Un autre phénomène intéressant à remarquer concerne le [a] français : alors qu'il est « hautement similaire » au [a] tchèque quand il est prononcé isolément et *similaire* en contexte labial, il est acoustiquement et perceptivement très peu *similaire* au [a] tchèque dans la suite [tat] et pas du tout *similaire* au [a] tchèque dans la suite [kak] où il est *similaire* aux [ɛ/ɛ:] tchèque.
- Alors que la voyelle [œ] prononcée isolément se trouve dans une zone F1/F2 inoccupées par les voyelles tchèques, elle est dans ce contexte perceptivement reliée en grande partie au [ɛ/ɛ:] tchèque. Ce même phénomène concerne la voyelle [ɔ] qui, quand elle est prononcée dans les suites [tɔt] et [kɔk], est acoustiquement *nouvelle* mais perceptivement *similaire* car majoritairement reliée au [o/o:] tchèque. Notons que malgré une plus grande similarité acoustique du [ɔ] français au [a] tchèque dans la suite [pɔp], cette voyelle reste dans cet environnement perceptivement reliée à la catégorie /o/ tchèque.

Pour résumer, nous constatons qu’une division des voyelles du français en *identiques*, *similaires* ou *nouvelles* par rapport aux voyelles tchèques est problématique. En effet, les voyelles acoustiquement *identiques* entre les deux langues n’existent pas. Les tests-t ont montré qu’au maximum un formant des voyelles proches est identique, les autres étant différents. Ensuite, aucune voyelle n’est *nouvelle* selon les trois critères dans tous les environnements phonétiques étudiés. Même si les voyelles [y] et [ø] sont acoustiquement toujours *nouvelles*, elles sont dans certains contextes perceptivement *similaires* à des voyelles maternelles. Ainsi, selon les trois critères de comparaison, toutes les dix voyelles orales du français actuel sont au moins dans quelques uns des contextes phonétiques *similaires* aux voyelles monophthongues du tchèque de Bohême. Il est néanmoins possible d’établir le degré de similarité entre les voyelles pour chaque contexte séparément en distinguant entre les voyelles “hautement similaires”, *similaires* et *nouvelles*.

En appliquant le SLM, nous prévoyons chez les apprenants de FLE de LM tchèque la possibilité d’acquérir authentiquement les voyelles “hautement similaires” [ɛ] en contextes [p, t, k] et [a] en isolation. De même, l’acquisition fidèle de la voyelle *nouvelle* [ø] en isolation et en contexte palato-vélaire et [y] en contexte dental est prévue. Les autres voyelles du français devraient être produites et perçues par la catégorie phonétique tchèque la plus proche.

Quoique l’effet de la graphie sur l’apprentissage d’une LE ne soit pas intégré dans les hypothèses et postulats de SLM, nous prévoyons, en accord avec la littérature exposée dans le chapitre 5 (section 5.6.1), une difficulté particulière due à la divergence des symboles phonétiques utilisés. Il s’agit de la voyelle [e] qui est dans tous les contextes *similaire* au [ɪ] tchèque, du [o] fermé isolé, acoustiquement *similaire* au [u] tchèque, du [a] *similaire* en contexte palato-vélaire au [ɛ:] long tchèque et du [ɔ] *similaire* au [a] tchèque en contexte labial.

Troisième partie : Voyelles orales du français des tchécophones par rapport à la « référence »

Dans la troisième partie (chapitres 12 – 14), nous nous intéressons aux limites effectives de la maîtrise phonétique des voyelles orales du français chez des futures enseignantes de FLE de LM tchèque. Nous comparons leur performance avec la « référence », représentée dans notre étude par des francophones natifs non-méridionaux.

12. Voyelles orales du français chez des futures enseignantes de FLE de LM tchèque

Au teint, on juge l'étoffe ; au bouquet, le vin ; à l'odeur, la fleur ; au langage, l'homme.

Charles Cahier

Résumé : Le **chapitre douze** s'intéresse aux limites de l'approximation de la cible des dix voyelles orales du français par dix futures enseignantes tchèques de FLE et à la fiabilité des prédictions établies sur la base du modèle SLM. La maîtrise en production chez les tchécophones est estimée à partir de :

- la comparaison des formants des voyelles en isolation et en syllabe finale pVp, tVt, kVk et RVR (4 répétitions*10 futures enseignantes) avec la « référence » (4 répétitions*10 Françaises natives non-méridionales). Les voyelles produites à moins d'un écart type de la « référence » sont jugées comme à priori authentiques, celles produites à plus de deux écarts types de la « référence » comme non maîtrisées,
- des résultats des tests d'identification par des Français natifs des voyelles isolées (100 réponses par voyelle : Voyelle isolée*10 locutrices tchécophones*10 auditeurs français natifs). Les tests permettent de confirmer ou infirmer l'interprétation des résultats acoustiques pour les voyelles isolées.

La maîtrise en perception est ensuite estimée à partir d'identifications par des futures enseignantes de FLE des voyelles isolées produites par une Française parisienne (80 réponses par voyelle : Voyelle isolée* 8 répétitions par une Française native*10 auditrices tchécophones).

Le résultat est en concordance avec les prédictions de SLM concernant la maîtrise des voyelles *similaires* aux voyelles thèques. Ces dernières sont en général produites et perçues au travers de la catégorie tchèque phonétiquement la plus proche et le degré d'authenticité de la voyelle du français est proportionnel au degré de similarité qu'elle entretient avec son équivalent tchèque. Ainsi dans les contextes où les voyelles du français sont "hautement similaires" à une voyelle tchèque, leur maîtrise par les tchécophones est généralement authentique. A l'inverse plus la voyelle du français est acoustiquement éloignée du correspondant tchèque le plus proche, moins sa cible est atteinte par les futures enseignantes, à l'exception du [i]. Ce résultat est en accord avec Flege (1987d) qui explique les limites de l'apprentissage phonétique des voyelles *similaires* par un mécanisme cognitif de *classification par équivalence*. Il est cependant nécessaire de souligner l'immense influence de la graphie qui prime en général sur la similarité phonétique. En effet malgré une grande similarité acoustico-perceptive du [e] français (noté généralement par le graphème <é>) avec le [i] bref tchèque (noté par le graphème <i>), cette voyelle est perçue et produite par les tchécophones sous l'influence de la graphie au travers de la catégorie tchèque [ɛ].

Le résultat concernant l'apprentissage des voyelles *nouvelles* est partiellement en concordance avec les hypothèses de SLM. De nouvelles catégories phonétiques pour des voyelles acoustiquement *nouvelles*, retranscrites avec des symboles phonétiques inexploités dans le système vocalique tchèque, peuvent être formées dans l'esprit des Tchécophones, comme le prévoit le SLM. En général plus leur similarité perceptive avec une voyelle tchèque est petite, plus la cible acoustique est atteinte. En revanche, les voyelles *nouvelles* selon les trois critères (donc même au plan perceptif) ne sont pas toujours maîtrisées avec authenticité, ce qui est contre les attentes de SLM

Afin de cibler les voyelles du français dont l'acquisition authentique par des tchécophones peut être limitée, nous avons utilisé le modèle d'apprentissage phonétique Speech Learning Model (SLM) de Flege (1995) qui distingue les sons *nouveaux*, *similaires* et *identiques* aux sons de la LM.

Selon ce modèle, l'authenticité de la perception et de la production des sons étrangers est fortement liée à la similarité que ces sons entretiennent avec les sons de la LM : les sons *identiques* et *nouveaux* peuvent être maîtrisés avec une grande précision alors qu'il est peu probable d'acquérir les sons *similaires* aux catégories maternelles authentiquement.

Les résultats de la deuxième partie de la thèse montrent que les similarités acoustique et perceptive des voyelles du français et du tchèque dépend largement du contexte phonétique dans lequel les voyelles se trouvent.

Des voyelles *identiques* entre les deux langues n'existent pas et les sons véritablement *nouveaux*, acoustiquement, perceptivement et « symboliquement » ne sont qu'au nombre de deux : le [y] et le [ø]. Ainsi, selon les trois critères, le [y] est *nouveau* en contexte coronal et le [ø] en contextes nul et palato-vélaire, comme exposé en conclusion partielle de la deuxième partie.

Les huit autres voyelles orales [i, e, ε, u, o, ɔ, a, œ] sont toutes perceptivement *similaires* à une ou plusieurs voyelles tchèques et la plupart sont *similaires* du point de vue de leurs formants (toutes mis à part le [œ] en isolation et le [ɔ] en contextes palato-vélaire et coronal), puis « symboliquement » *similaires*, mis à part le [œ] qui est un symbole inexploité dans le système tchèque. Le degré et le type de similarité de ces voyelles peuvent changer d'un contexte à l'autre et le SLM ne précise pas comment le type de similarité influe sur la possibilité d'acquisition plus ou moins authentique.

En examinant si une plus grande similarité d'une voyelle non-native avec une voyelle native facilite ou au contraire complique un apprentissage fidèle et si son acquisition dépend du contexte phonétique (c'est-à-dire qu'une voyelle peut être maîtrisée de manière native dans un contexte phonétique et elle peut ne pas l'être dans un autre contexte), nous nous sommes posée les cinq questions suivantes :

1. Les voyelles véritablement *nouvelles*, c'est-à-dire [y] en contexte coronal et [ø] en isolation et en contexte palato-vélaire, sont-elles acquises de manière authentique ?
2. La plupart des voyelles françaises étant *similaires* aux voyelles tchèques, existe-il un lien entre le type/ degré de leur similarité avec les voyelles de la LM et le rapprochement de la « référence » de la LE ? Autrement dit, les voyelles du français « hautement similaires » (avec similarité +++) aux voyelles tchèques selon les trois critères (« symbolique », acoustique et perceptif) sont-elles acquises de manière plus authentique que les voyelles moins *similaires* (similarité ++ ou similarité +) ou que les voyelles *similaires* seulement selon certains des critères de comparaison ?
3. Le rapprochement de la « référence » diffère-t-il selon le contexte dans lequel la voyelle se trouve ? Autrement dit, est-il possible que la voyelle soit acquise de manière native dans certains contextes seulement ?
4. Y a-t-il des catégories vocaliques non maîtrisées par les futures enseignantes ? Si oui, quel est le type/ degré de similarité de ces catégories non maîtrisées avec les voyelles de la LM ? Une voyelle non maîtrisée en production se réaliserait à plus de deux écarts types de la même voyelle produite par un natif et serait perceptivement interprétée par les Français natifs en termes d'une autre catégorie, ou bien de plusieurs catégories ou encore non assimilée.
5. Enfin, qu'en est-il des voyelles en contexte uvulaire dont la similarité n'a pas pu être définie du fait de l'absence de la consonne uvulaire dans le système consonantique tchèque ?

Pour répondre aux questions, nous avons comparé les valeurs formantiques moyennes (voir les pages 85 et 86 des annexes) des voyelles du français produites par dix Tchèques, futures enseignantes de FLE, avec celles des voyelles produites par dix Françaises natives (productions qui représentent la « référence » dans cette thèse).

Chaque locutrice a répété la voyelle isolée et la voyelle insérée dans des logatomes du type CVCVCVC quatre fois, comme précisé dans la méthodologie en 6.2 et récapitulé dans le Tableau 91. Puisque dans l'enseignement de la prononciation du français on préconise le travail sur la qualité vocalique de la syllabe finale du mot prosodique qui est une position accentuée forte (Wioland, 2005), seules les voyelles isolées et les voyelles de la syllabe finale des logatomes trisyllabiques ont été examinées. En effet, comme l'a écrit l'auteur à la page 29 de son ouvrage La vie sociale des sons du français : « *Quelle que soit la stratégie didactique employée, c'est la dernière syllabe prononcée du mot phonétique qui doit faire l'objet de toutes les attentions.* ».

Corpus	10 voyelles orales du français [i, e, ε, y, ø, œ, u, o, ɔ, a] en isolation et insérées dans des logatomes pVpVpVp, tVtVtVt, kVkVkVk, RVRVRVR d'une phrase cadre, exemple : « Le mot papapape peut bien coller. », répétées 4 fois
Locutrices natives	10 Françaises natives non-méridionales ayant grandi ou longtemps séjourné en Ile-de-France
Locutrices non-natives	10 Tchèques de la région de Bohême, futures enseignantes de FLE
Calcul de la moyenne formantique	A partir de la valeur relevée à un tiers, à la moitié et à deux tiers de la durée des voyelles isolées et à la moitié de la durée de chaque voyelle finale des logatomes trissyllabiques (10 locutrices* 4 répétitions)

Tableau 91 : Description du corpus et du calcul des moyennes formantiques des voyelles françaises produites en isolation et en contexte par des natives du français (la « référence ») et les non-natives

Afin de déterminer le niveau de maîtrise dans la production des sons vocaliques, nous avons observé les distances acoustiques entre les formants moyens F1, F2 et F3 des voyelles produites par les Tchèques (moyennes données à la page 87 des annexes) par rapport à celles des natives (les valeurs formantiques moyennes des voyelles prononcées en syllabe finale des logatomes se trouvent en annexe du chapitre 12, à la page 88 des annexes). La distance acoustique est exprimée en nombre d'écarts types, comme expliqué dans le chapitre 6 sur la méthodologie et cela pour tous les contextes séparément. Nous avons ensuite fait écouter les voyelles isolées des futures enseignantes tchèques aux Français natifs afin de vérifier leur identification. Notons que la perception par des Français natifs des voyelles produites par les non-natives en contexte consonantique n'a pas été vérifiée et ce point fera l'objet de nos études futures. Le Tableau 92 récapitule la construction et les conditions de passation du test d'identification des 13 voyelles isolées du français (y compris les nasales), produites par les Tchécophones (une série) et perçues par 10 Français natifs. Il est à noter que nous n'avons utilisé que les réponses concernant les dix voyelles orales.

Afin de vérifier le niveau de maîtrise des catégories perceptives des voyelles du français par les futures enseignantes tchèques, nous leur avons proposé un test d'identification des voyelles isolées produites par une Française parisienne. Notons que la perception des voyelles produites en contexte consonantique n'a de nouveau pas été vérifiée. Le Tableau 93 de la page 242 récapitule la construction et les conditions de passation du test d'identification des voyelles isolées du français, produites par les Français natifs (une série) et perçues par 10 futures enseignantes tchèques.

Type de test	D'identification
Stimuli	13 voyelles isolées du français [i, e, ε, y, ø, œ, u, o, ɔ, a, ã, ě, ǐ], une série par locutrice
Consigne	Vous allez entendre des voyelles. Identifiez ces voyelles en cliquant sur le bon mot. Indiquez si vous êtes sûr(e) de votre choix sur une échelle de 1 à 5 où 5 correspond à j'en suis complètement sûr(e) 1 à je n'en suis pas sûr(e) du tout
Auditeurs	10 Français natifs
Locuteurs	10 Tchèques ayant un niveau avancé en français
Nb total de stimuli	139
Nb de stimuli d'entraînement	9
Nb de stimuli distracteurs	0
Echelle de justesse	1 – 5 1 pas sûr(e) du tout 3 moyennement sûr(e) 5 sûr(e)

Tableau 92 : Information sur la construction et les conditions de passation du test d'identification des voyelles isolées du français produites par dix tchécophones et identifiées par dix Français natifs non-méridionaux

Comme le fait apparaître le Tableau 93 de la page 242, le test d'identification contient 111 stimuli : treize voyelles isolées du français [i, e, ε, y, ø, œ, u, o, ɔ, a, ã, ě, ǐ], répétées huit fois par une Française, qui est née et a vécu toute sa vie à Paris, et sept stimuli d'entraînement. Dans cette étude, nous n'avons utilisé que les réponses concernant les dix voyelles orales.

Les résultats des comparaisons acoustiques et des tests perceptifs sont systématiquement interprétés en lien avec le type/ degré de similarités perceptive et acoustique que chacune des voyelles françaises entretient avec les voyelles tchèques. Ainsi nous examinons d'abord l'acquisition des voyelles *nouvelles*, puis des voyelles *similaires* aux voyelles tchèques en les classant en plusieurs groupes, selon le type et le degré de similarité qu'elles entretiennent avec les voyelles de la LM. Notons que seule l'acquisition des voyelles produites en contexte uvulaire n'a pas pu être interprétée en termes de similarité acoustico-perceptive du fait du manque de la consonne uvulaire /ʁ/ dans le système tchèque.

Type de test	D'identification
Stimuli	13 voyelles isolées du français [i, e, ε, y, ø, œ, u, o, ɔ, a, ã, ê, ÿ], 8 séries
Consigne	Vous allez entendre des voyelles. Identifiez ces voyelles en cliquant sur le bon mot. Indiquez si vous êtes sûr(e) de votre choix sur une échelle de 1 à 5 où 5 correspond à j'en suis complètement sûr(e) et 1 à je n'en suis pas sûr(e) du tout
Auditeurs	10 Tchèques ayant un niveau avancé en français
Locuteurs	1 Française parisienne
Nb total de stimuli	111
Nb de stimuli d'entraînement	7
Nb de stimuli distracteurs	0
Echelle de justesse	1 – 5 1 pas sûr(e) du tout 3 moyennement sûr(e) 5 sûr(e)

Tableau 93 : Information sur la construction et les conditions de passation du test d'identification des voyelles isolées du français produites par une Française parisienne et identifiées par dix Tchèques, futures enseignantes de FLE

12.1 Voyelles nouvelles

Le SLM prévoit la possibilité d'acquérir les voyelles *nouvelles* de façon native. Comme noté auparavant, seules deux voyelles sont véritablement *nouvelles*, selon les critères acoustique, perceptif et "symbolique" : le [y] en contexte coronal et le [ø] en isolation et en contexte palato-vélaire.

Ainsi, en concordance avec le SLM, nous faisons l'hypothèse que la voyelle [y] en syllabe finale [tyt] ainsi que la voyelle [ø] en isolation et en syllabe finale [køk] peuvent être produites par les dix Tchèques, futures enseignantes de FLE, authentiquement, c'est-à-dire avec les formants F1, F2 et F3 à moins d'un écart type de ceux des [y, ø] prononcés par les dix natives (productions que nous considérons comme étant la « référence »). La perception authentique par les futures enseignantes de FLE n'a été vérifiée que pour la voyelle isolée [ø].

Afin de vérifier si les Tchèques possèdent la catégorie perceptive pour [ø] isolé, nous leur avons fait passer un test d'identification de toutes les voyelles du français, décrit dans le Tableau 93. Le résultat à propos de l'identification du [ø], produit huit fois en isolation par une Française parisienne native se trouve dans le Tableau 94. Nous y remarquons que dans 84 % des réponses, les stimuli ont été "correctement"¹⁹ identifiés, avec une note de 4,1 sur 5 (soit un *fit index* de 3,4) ce qui

¹⁹ « correctement » correspond à une identification en termes de la catégorie que le locuteur avait intention de produire

montre que [ø] isolé n'est pas toujours perçu authentiquement. En effet, dans 13 % des réponses, il a été confondu avec /œ/ et dans 1 % des réponses avec /ɛ, o, ẽ/.

Stimulus	Voyelles identifiées	Pourcentage de l'identification	Echelle de confiance	Fit index
ø	ø	84	4,1	3,4
	œ	13	2,5	0,3
	ɛ	1	4	0,04
	o	1	4	0,04
	ẽ	1	2	0,02

Tableau 94 : Identification par 10 Tchèques, futures enseignantes de FLE, du [ø] produit en isolation par une Française parisienne, répétées 8 fois. Le stimulus est en colonne 1, les réponses en colonne 2, le pourcentage de l'identification en colonne 3, la note moyenne en colonne 4 et le *fit index*, calculé selon Guion *et al.* (2000), en colonne 5

Le Tableau 95 renseigne ensuite sur les valeurs formantiques F1, F2 et F3 (avec les écarts types entre parenthèses) de la voyelle [y] en contexte coronal et [ø] en isolation et en contexte palato-vélaire, produites par dix Françaises natives (Nat) et dix Tchèques, locutrices non-natives (Non-nat). Les voyelles en contexte sont celles de la syllabe finale des logatomes CVCVCVC.

Voyelles/formants	F1		F2		F3	
	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat
y (t)	286 (29, 10,1%)	274 (33, 12%)	2095 (164, 7,8%)	2109 (135, 6,4%)	2697 (180, 6,7%)	2782 (186, 6,7%)
ø (0)	400 (35, 8,6%)	416 (59, 14,2%)	1571 (127, 8,1%)	1566 (173, 11%)	2645 (173, 6,5%)	2718 (88, 3,2%)
ø (k)	408 (46, 11,3%)	419 (58, 13,8%)	1749 (105, 6%)	1584 (256, 16,2%)	2486 (174, 7%)	2588 (187, 7,2%)

Tableau 95 : Moyennes et écarts types de F1, F2 et F3 (en Hertz) de [y] en contexte tVt et de [ø] en isolation (0) et en contexte kVk, produits par 10 Françaises natives (Nat) et 10 Tchèques (Non-nat), 4 répétitions. Les cases avec les valeurs à plus d'un écart type de la « référence » sont grisées

En étudiant le Tableau 95, nous constatons que les voyelles [y] en contexte coronal ainsi que [ø] en isolation sont produites de manière à priori authentique (car leurs formants sont réalisées à moins d'un écart type de la « référence »). En revanche, le deuxième formant de la voyelle [ø] en contexte palato-vélaire, produite par les tchécophones, se situe à plus d'un écart type de la « référence ». Les triangles vocaliques F1/F2, F2/F3 (en Bark) de la Figure 100 de la page 244 illustrent la position acoustique du [ø] final en contexte palato-vélaire produit par les Tchèques (en pointillé) par rapport à la « référence » (en trait plein). L'écart type affiché est de un.



Figure 100 : Triangles vocaliques F1/F2, F2/F3 (en Bark) des voyelles orales du français produites en syllabe finale du logatome kVkvkVkvk (4 répétitions) par dix Françaises (trait plein) et dix Tchèques (trait pointillé). Les ellipses de dispersion sont tracées à un écart type. Notons que le [ø] des Tchèques se trouve à plus d'un écart type de la « référence »

Le [ø] prononcé en syllabe finale [køk] par les Tchèques diffère ainsi de la « référence » au niveau du deuxième formant. Cependant ce résultat ne doit pas être considéré comme de l'échec dans l'atteinte de l'authenticité car selon la loi de position qui régit de plus en plus la distribution des voyelles moyennes en français standard, le [ø] tend à apparaître en syllabe finale ouverte plutôt qu'en syllabe fermée. Ce point est traité plus en détail dans les chapitres 13 et 14 portant sur les voyelles moyennes.

Notons également chez les non-natives un écart type élevé de la moyenne du F2 du [ø] produit en syllabe finale [køk] qui est de 256 Hz, ce qui correspond à une variation de 16,2 % (voir Tableau 95). Cette donnée montre une forte variation dans la production du [ø] au sein du groupe des non-natives. Le Tableau 96 affiche les valeurs moyennes du deuxième formant calculées pour chaque locutrice non-native séparément.

Non-nat	F2 de [ø] dans [køk] (en Hz)
T1	1629
T2	1309
T3	1197
T4	2026
T5	1562
T6	1099
T7	1595
T8	1430
T9	1783
T10	1835

Tableau 96 : Le F2 moyen (en Hz) de [ø] dans [køk] chez les 10 non-natives (non-nat) séparément (moyenne de 4 répétitions)

Le Tableau 96 montre que la variation de la valeur du deuxième formant selon la locutrice est importante : elle varie en moyenne de 1009 à 2026 Hz. Ainsi le F2 de [ø] dans [køk] se trouve à moins d'un écart type de la « référence » dans la production des locutrices T9 et T10 uniquement (Non-nat, colonne 1).

Le résultat du test d'identification des voyelles isolées montre que le [ø] produit en isolation par les locutrices non-natives est identifié par les natifs en tant que /ø/ dans 78 % des réponses, avec une note moyenne de 4,7. Le *fit index* de l'identification du [ø] est alors de 3,7. De cette manière, malgré le fait que la voyelle [ø] soit *nouvelle* en isolation et devrait donc être maîtrisée authentiquement, elle est moyennement identifiable par les Français natifs.

Nos données ne sont alors que partiellement en concordance avec les prédictions du SLM à propos de la possibilité de l'apprentissage des sons *nouveaux*. Ces derniers peuvent être appris, comme le prédit SLM mais leur authenticité n'est pas toujours atteinte par les futures enseignantes de FLE. Alors que le [y] en contexte coronal est produit fidèlement, malgré un plus grand écartement des formants F3/F2 par rapport à la « référence », et donc un caractère focal plus faible, le [ø] de la syllabe finale du logatome [køkøkøk] s'écarte de la « référence » (ce qui n'est pas nécessairement de l'échec si l'on suit la tendance du français actuel à distribuer la qualité des voyelles moyennes en fonction de la structure syllabique) et le [ø] isolé, malgré une réalisation acoustique proche de la « référence » est moyennement identifiable par les natifs. De même, la perception du [ø] isolé par les Tchèques n'est pas toujours authentique. Notons que l'identification des voyelles en contexte (que ce soit par des Français natifs ou par des Tchèques) n'a pas été vérifiée et fera l'objet de nos études futures.

12.2 Voyelles similaires

Comme nous l'avons constaté dans la deuxième partie de la thèse, toutes les voyelles du français peuvent être *similaires* aux voyelles tchèques dans certains contextes phonétiques, du moins selon l'un des trois critères. Le Tableau 97 récapitule les voyelles françaises *similaires* aux voyelles tchèques selon l'environnement phonétique en les regroupant selon le type de similarité. « A » indique que la voyelle du français entretient une similarité acoustique et « P » une similarité perceptive avec une voyelle tchèque. « S » indique une similarité au niveau des symboles phonétiques (par exemple [i] français est *similaire* au [i:] tchèque), (S) entre parenthèses au contraire indique une divergence des symboles phonétiques (par exemple [e] français est acoustiquement et perceptivement *similaire* au [ɪ] tchèque). Enfin, une voyelle française retranscrite avec un symbole phonétique non utilisé dans le système vocalique tchèque (absence de S) peut être *similaire* perceptivement et parfois acoustiquement à une voyelle tchèque.

Groupe	Voyelle	0	p	t	k
1	i	S-A-P	S-A-P	S-A-P	S-A-P
	ε	S-A-P	S-A-P	S-A-P	S-A-P
	u	S-A-P	S-A-P	S-A-P	S-A-P
2	a	S-A-P	S-A-P	(S)-A-P	(S)-A-P
	e	(S)-A-P	(S)-A-P	(S)-A-P	(S)-A-P
	o	(S)-A-P	(S)-A-P	(S)-A-P	(S)-A-P
3	ɔ	S-A-P	(S)-A-P	S-P	S-P
4	œ	P	A-P	A-P	P
5	y	P	P	--	P
	ø	--	P	P	--

Tableau 97 : Similarité “symbolique” (S), acoustique (A) et perceptive (P) des dix voyelles françaises avec les voyelles tchèques selon le contexte phonétique. Divergence “symbolique” marquée entre parenthèses : malgré l'existence du symbole [e, o] en tchèque, le [e, o] français se réalise et/ou est identifié comme respectivement [ɪ, u/u:] tchèques

Nous remarquons dans le Tableau 97 que les voyelles qui sont *similaires* à des voyelles tchèques selon les trois critères et dans tous les contextes sont le [i, ε, u] qui constituent ainsi le premier groupe vocalique.

Le deuxième groupe correspond aux voyelles [a, e, o] qui entretiennent une similarité phonétique avec une voyelle tchèque retranscrite avec un symbole différent (même si le symbole de la voyelle française est également utilisée dans le système tchèque mais pour noter un timbre différent). Alors que le [a] du français est acoustiquement et perceptivement *similaire* au [a] tchèque en contextes nul et labial, elle est acoustiquement et perceptivement *similaire* au [ε/ε:] tchèque en contextes coronal et palato-vélaire. La voyelle [e] est acoustiquement et perceptivement *similaire* au [ɪ] tchèque dans tous les contextes et enfin la voyelle [o] est acoustiquement *similaire* au [o/o:] tchèque en contextes [p, t, k] et au [u/u:] tchèque en isolation et elle est perceptivement *similaire* au [u/u:] tchèque dans tous les

contextes. Nous remarquons ainsi une divergence de symboles utilisés pour deux sons acoustiquement et perceptivement proches, du moins dans certains des contextes phonétiques étudiés.

La voyelle [ɔ] du groupe 3 entretient une similarité plus limitée avec les voyelles tchèques que les voyelles du groupe 1 ou 2. Alors qu'elle est perceptivement *similaire* dans tous les contextes au [o/o:] tchèque, elle rencontre une similarité acoustique avec [o/o:] tchèque en isolation et avec [a/a:] tchèque en contexte labial. Dans les autres contextes [ɔ] est une voyelle acoustiquement nouvelle. Notons de nouveau que lorsque [ɔ] se trouve en isolation, il est *similaire* à une voyelle tchèque retranscrite par un autre symbole phonétique (qui est le [a]).

La voyelle [œ] du groupe 4 est contrairement aux voyelles du groupe 1, 2 et 3 “symboliquement” *nouvelle*, ce qui veut dire que ce symbole phonétique n'est pas utilisé pour décrire le système vocalique du tchèque. En revanche, [œ] est *similaire* au [ɛ/ɛ:] tchèque perceptivement (dans tous les contextes) et acoustiquement (en contextes coronal et palato-vélaire). En isolation et en contexte labial, [œ] est une voyelle acoustiquement *nouvelle*.

Enfin, le groupe 5 est constitué de voyelles [y, ø]. Comme dans le cas de [œ], il s'agit de symboles inexploités pour représenter les voyelles du tchèque. Mais à l'inverse du [œ], les voyelles [y, ø] sont acoustiquement *nouvelles* dans tous les contextes phonétiques étudiés. Leur similarité perceptive se limite ensuite aux environnements nul, labial et palato-vélaire pour [y] qui y est *similaire* au [ɪ] ou [i:] tchèque et aux environnements labial et coronal pour [ø] qui y est *similaire* respectivement au [ɛ/ɛ:] et [u/u:] tchèques. Nous remarquons que ces voyelles entretiennent ainsi une similarité moindre avec les catégories tchèques que les voyelles des autres groupes au point qu'on puisse les considérer dans certains contextes comme étant *nouvelles*.

En divisant les voyelles du français en cinq groupes selon la similarité qu'elles entretiennent avec les voyelles tchèques sur le plan “symbolique” et/ou perceptif et/ou acoustique, nous pouvons étudier la possibilité de se rapprocher de la « référence » en fonction du type de similarité. Notons néanmoins que ce regroupement est schématique car à l'intérieur de chaque groupe, le degré et le type de similarité effective est toujours fonction du contexte phonétique où la voyelle est insérée. De cette façon, alors que la voyelle [a] est par exemple classée dans le groupe 2 car elle peut être interprétée en termes d'une voyelle tchèque retranscrite par un autre symbole de l'API (qui est le [ɛ/ɛ:], en contextes coronal et palato-vélaire), elle pourrait faire partie du groupe 1 lors de sa production en contextes nul ou labial (où elle est *similaire* selon les trois critères étudiés au [a/a:] tchèque).

12.2.1 Groupe 1 : Voyelles [i, ɛ, u], similarité maximale

12.2.1.1 Voyelle [i]

Les résultats de la deuxième partie et notamment le Tableau 90 de la page 231 montrent que [i] français est maximale *similaire* au [i:] tchèque en contexte labial. En isolation, une plus grande distance acoustique sépare les deux voyelles *similaires* et en contextes coronal et palato-vélaire, elles s'éloignent davantage perceptivement.

Pour vérifier si les tchécophones possèdent une catégorie perceptive pour le [i] français isolé, nous leur avons fait passer un test d'identification dont la construction et les conditions de passation sont décrites dans le Tableau 93 de la page 242. Le résultat révèle que dans 99 % des réponses, les stimuli ont été “correctement” identifiés avec une note moyenne de 4,8 sur 5 (soit un *fit index* de 4,7) et nous en concluons que les locutrices non-native perçoivent le [i] français authentiquement.

Concernant la production, la Figure 101 de la page 248 montre que les formants F1, F2, F3 et F4 (en Bark) du [i] produit par les dix Tchèques (marqué par un triangle) en isolation (traits verticaux) et en contextes [p, t, k, ʁ] se trouvent à un écart type ou moins de la « référence » (marquée par une

croix), calculée à partir de productions de dix Françaises natives. Le [i] est alors réalisé à priori fidèlement par les non-natives dans tous les contextes. Notons que seul le F1 de [i] de la syllabe médiane du logatome [kikikik] se réalise à plus d'un écart type de la « référence » ce qui n'attire pas notre attention car il s'agit d'une position prosodiquement faible qui est moins importante dans l'apprentissage phonétique du FLE (Wioland, 2005).

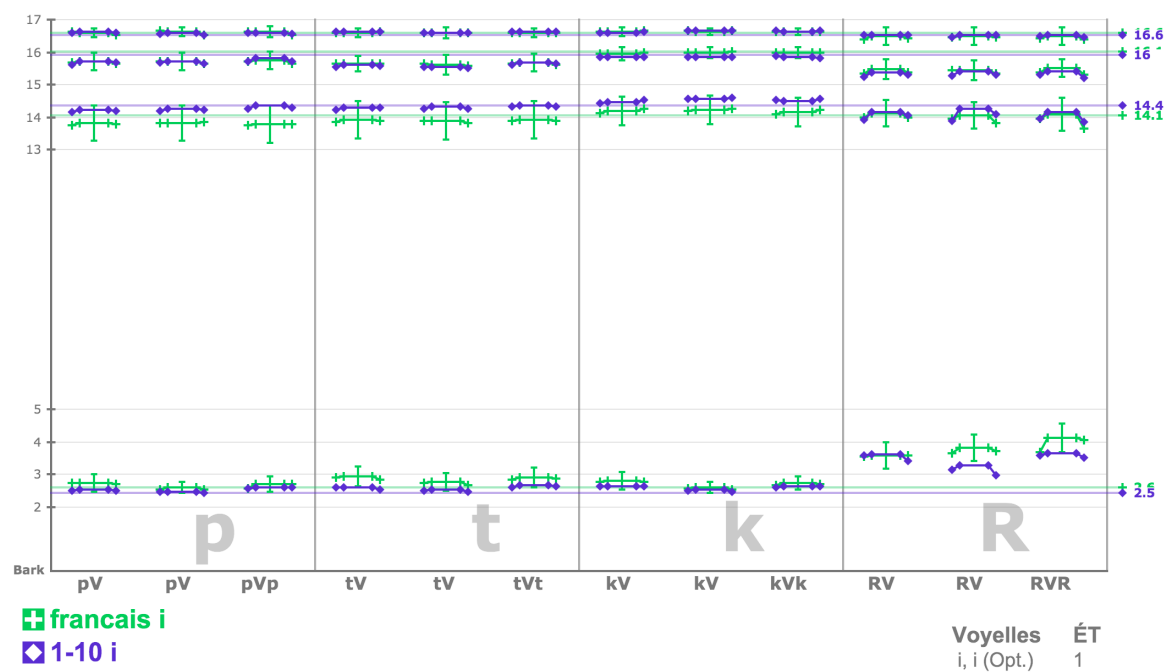


Figure 101 : Valeurs moyennes de F1, F2, F3 et F4 (en Bark) du [i] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un losange) et relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1

Nous avons ensuite étudié le rapprochement des troisième et quatrième formants du [i] qui est maximal lorsque la voyelle est prononcée en isolation et moindre lorsqu'elle est insérée en contexte phonétique. La distance acoustique F4-F3 (en Hz) du [i] produit par les Tchèques et par les Françaises natives est indiquée pour comparaison dans le Tableau 98.

	0		p		t		k		R	
F4-F3	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat
i	631	652	863	852	905	976	683	790	943	1023

Tableau 98 : Distance acoustique F4-F3 du [i] prononcé par 10 Françaises natives (Nat) et 10 Tchèques (Non-nat) en isolation (0) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], 4 répétitions par locutrice

Le Tableau 98 montre que le [i] prononcé par les Tchèques se caractérise par une distance acoustique F4-F3 supérieure à celle du [i] produit par les natives, mis à part en contexte labial. Cet écartement formantique est important en contextes [t, k, ʁ] (+ 70 Hz par rapport à la « référence »). Ainsi, le caractère focal du [i] en contextes coronal, palato-vélaire (où [i] est perceptivement moins *similaire* au [i:] tchèque qu'en isolation et en contexte labial) et uvulaire est affaibli. En interprétant ce résultat en lien avec les résultats à propos du degré de similarité entre le [i] français et le [i:] tchèque, nous constatons qu'une plus grande similarité perceptive entre le [i] français et le [i:] tchèque favorise sa production authentique par les Tchèques.

L'identification du [i] isolé par dix Français natifs confirme que cette voyelle avait été produite authentiquement car les stimuli ont été “correctement” identifiés dans 100 % des réponses avec une note moyenne de 4,9 sur 5 ; soit un *fit index* de 4,9.

Le résultat faisant apparaître la possibilité d'atteindre par les futures enseignantes l'authenticité du [i] français isolé, malgré sa relativement petite similarité acoustico-perceptive avec le [i:] tchèque, est en accord avec l'étude de Flege (1997) qui révèle une facilité à produire et percevoir les voyelles hautes de LE.

12.2.1.2 Voyelle [ɛ]

La voyelle du français [ɛ], comme précisé dans le Tableau 90 de la page 231, est « hautement similaire » perceptivement et acoustiquement au [ɛ/ɛ:] tchèque en contextes [p, t, k] et acoustiquement moins *similaire* lors qu'elle est prononcée en isolation (même si sa similarité perceptive avec le [ɛ/ɛ:] tchèque reste élevée). Qu'en-est-il alors de sa maîtrise par les Tchèques ?

Dans le but d'examiner si les non-natives possèdent une catégorie perceptive pour [ɛ] français en isolation, nous leur avons proposé un test d'identification décrit dans le Tableau 93. Le résultat montre que dans seulement 61 % des réponses, les stimuli ont été “correctement” identifiés comme /ɛ/ et ceci avec note moyenne de 3,8 sur 5 (soit un *fit index* de 2,3). Les autres stimuli ont été majoritairement confondus, comme nous pouvons le voir dans le Tableau 99, avec /e/ (dans 28 % des réponses, avec une note de 4 sur 5), puis avec /œ/ (dans 10 % des réponses, avec une note de 3,2 sur 5) et enfin avec /ø/ (dans 1 % des réponses seulement, avec une note de 1 sur 5). Nous pouvons en conclure que la catégorie perceptive pour [ɛ] français en isolation n'est pas formée chez les Tchèques.

Stimulus	Réponse	%identif ication	Note	<i>Fit index</i>
ɛ	ɛ	61	3,8	2,3
	e	28	4	1,1
	œ	10	3,2	0,3
	ø	1	1	0,01

Tableau 99 : Identification par 10 Tchèques, futures enseignantes de FLE, du [ɛ] produit en isolation par une Française parisienne, répétées 8 fois. Le stimulus est en colonne 1, les réponses en colonne 2, le pourcentage de l'identification en colonne 3, la note moyenne en colonne 4 et le *fit index*, calculé selon Guion *et al.* (2000), en colonne 5

Au sujet de la production, nous avons comparé les caractéristiques acoustiques du [ɛ] produit par les non-natives par rapport à la “référence”. Les valeurs formantiques moyennes calculées à partir de 10 locutrices et 4 répétitions de la Figure 102 montrent que les formants F1, F2 et F3 (en Bark) du [ɛ] produit par les dix Tchèques (marqué par un triangle) se trouvent à moins d'un écart type de la “référence” (marquée par une croix) dans tous les contextes, mis à part pour le F1 du [ɛ] de la syllabe initiale du logatome [tɛtɛtɛ]. Cette différence a de nouveau moins d'importance dans l'enseignement du français aux étrangers car il s'agit d'une position prosodique faible.

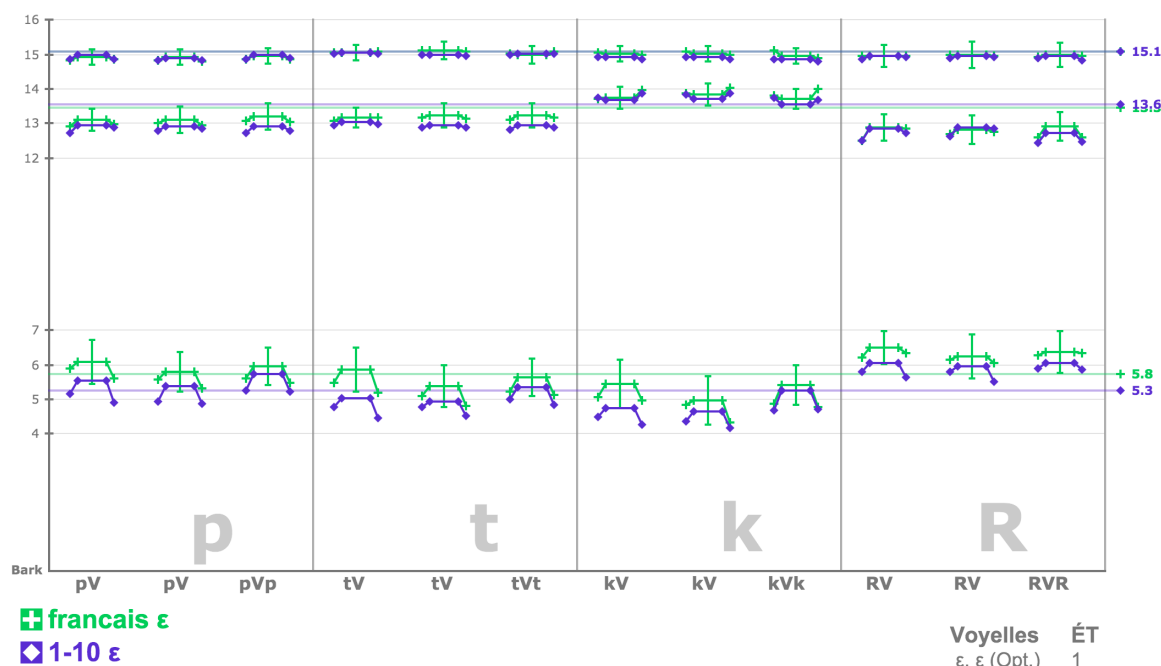


Figure 102 : Valeurs moyennes de F1, F2, F3 (en Bark) du [ε] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Français (marquées par une croix) et 10 Tchèques (marquées par un losange) et relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1

La voyelle [ε], « hautement similaire » dans la plupart des contextes aux [ε/ε:] tchèques, est produite authentiquement dans tous les contextes étudiés. L'identification du [ε] isolé par les Français natifs confirme sa production authentique : dans 92 % des réponses, les stimuli ont été « correctement » identifiés avec une note moyenne de 4,5 sur 5, soit un *fit index* de 4,1. Ce résultat est en accord avec Bohn and Flege (1990) considérant que les sons « hautement similaires » (pour ne pas dire *identiques*) peuvent être maîtrisés avec une grande précision. En revanche, le lien de causalité entre perception et production est remis en question car quoique les tchécophones produisent le [ε] isolé authentiquement, elles le confondent avec /e/.

12.2.1.3 Voyelle [u]

Enfin, comme le récapitule le Tableau 90 de la page 231, la voyelle française [u] est acoustiquement *similaire* dans tous les contextes aux [u/u:] tchèques, avec les formants à 1,5 écarts types ou moins de ceux des [u/u:] tchèques. La similarité perceptive avec les [u/u:] tchèques est grande dans tous les contextes (*fit index* de 4,4 ou plus sur 5) mis à part en isolation où le *fit index* d'identification interlangue est de 3,9 sur 5.

Concernant la perception des Tchèques, il s'avère qu'elles possèdent une catégorie perceptive pour le [u] français isolé : dans le test d'identification décrit dans le Tableau 93 de la page 242, elles l'ont « correctement » identifié dans 96 % des réponses avec une note moyenne de 4,6 sur 5, soit un *fit index* de 4,4.

La Figure 103 compare ensuite la réalisation acoustique en termes de F1 et F2 (en Bark) du [u] prononcé par les dix Tchèques (marqué par un triangle) par rapport à la « référence » (marquée par une croix). L'écart type affiché est de un.

12. Voyelles orales du français chez des futures enseignantes de FLE de LM tchèque

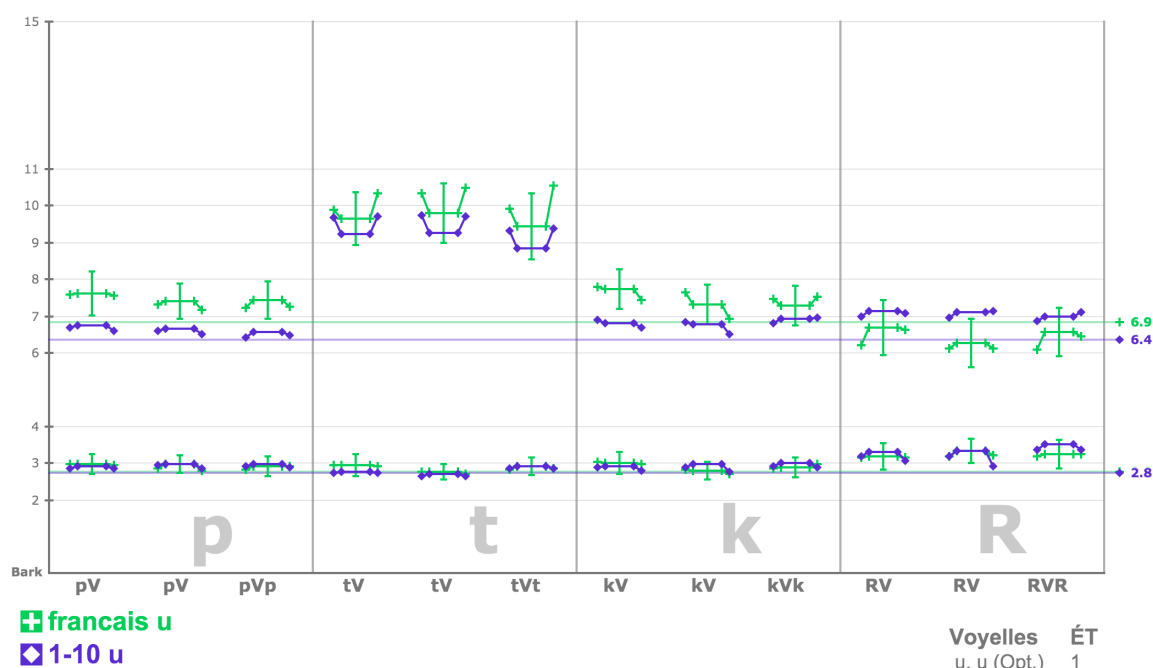


Figure 103 : Valeurs moyennes de F1, F2 (en Bark) du [u] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un losange) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1

Nous remarquons que dans tous les contextes et toutes les syllabes, le F1 du [u] prononcé par les dix Tchèques se trouve à moins d'un écart type de la "référence". En revanche, le F2 diverge davantage : il se trouve alors à moins de 1,5 écarts types en contextes labial (toutes les syllabes), palato-vélaire (en syllabe initiale) et uvulaire (en syllabe médiane). Dans les autres contextes et syllabes, il se réalise à moins d'un écart type de la "référence". Ainsi, mis à part l'environnement labial, le [u] de la syllabe finale et le [u] isolé est réalisé authentiquement par les non-natives.

Le rapprochement des premier et deuxième formants pour le [u] produit par les non-natives est comparé à celui de la "référence" dans le Tableau 100.

	0		p		t		k		R	
F2-F1	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat
u	495	424	551	438	918	840	551	462	383	450

Tableau 100 : Distance acoustique F2-F1 du [u] prononcé par 10 Françaises natives (Nat) et 10 Tchèques (Non-nat) en isolation (0) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], 4 répétitions par locutrice

Nous remarquons dans le Tableau 100 que les tchécophones rapprochent les deux premiers formants du [u] français davantage que les Françaises natives, mis à part en contexte uvulaire. Ce rapprochement formantique est alors semblable au rapprochement formantique du [u] du tchèque, qui a un caractère focal fort.

L'identification par des Français natifs du [u] isolé produit par les tchécophones a confirmé sa réalisation fidèle. Les auditeurs l'ont "correctement" identifié dans 96 % des réponses avec une note moyenne de 4,6 sur 5, soit un *fit index* de 4,4.

Globalement les voyelles [i, ε, u], *similaires* aux voyelles tchèques selon les trois critères de comparaison, sont réalisées par les Tchèques avec des valeurs formantiques se trouvant à moins d'un écart type de la "référence", du moins en syllabe finale, qui doit faire l'objet de l'apprentissage phonétique du FLE (Wioland, 2005). Seule la voyelle [u] en contexte labial est produite avec un F2 qui se trouve à moins de deux écarts types de la "référence". Les trois voyelles isolées sont également produites avec les formants à un écart type de la "référence", mis à part le [ε] qui est produit avec les formants à moins de 1,5 écarts types de la "référence", et elles sont identifiées par les Français natifs avec facilité ([ε] néanmoins un peu moins facilement que [i] et [u]). Ces voyelles sont alors produites par les futures enseignantes de FLE avec précision.

En revanche, la catégorie perceptive n'est formée que pour les voyelles isolées [i] et [u]. En effet, la voyelle antérieure mi-ouverte [ε] est confondue par les non-natives avec le /e/ mi-fermé.

Le Tableau 101 récapitule le niveau d'authenticité de [i, ε, u] exprimé en production par la distance acoustique par rapport à la "référence" et le taux d'identification (valeur du *fit index*) par les Français natifs et en perception par le taux d'identification par les non-natives.

Voyelle	[i]					[ε]					[u]				
Similarité avec une voyelle tchèque	S-A-P					S-A-P					S-A-P				
Contexte	0	p	t	k	R	0	p	t	k	R	0	p	t	k	R
Distance acoustique max. (en nb d'écarts types) des moyennes de F1, F2, (F3, F4) de la "référence"	1	1	1	1	1	<1,5	1	1	1	1	1	< 2	1	1	1
Le <i>fit index</i> de l'identification par les natifs (sur 5)	4,9					4,1					4,9				
Le <i>fit index</i> de l'identification par les non-natifs (sur 5)	4,7					2,3					4,4				

Tableau 101 : Production et perception par 10 Tchèques des voyelles françaises [i, ε, u] du groupe 1 se caractérisant par une similarité "symbolique" (S), acoustique (A) et perceptive (P) avec les voyelles tchèques. Le niveau en production est estimé selon la distance des formants par rapport à la "référence" (exprimée en nombre d'écarts types) et le *fit index* de l'identification par 10 Français natifs de chaque voyelle produite par 10 Tchèques en isolation (100 réponses par voyelle). Le niveau de perception est estimé par le *fit index* de l'identification par 10 Tchèques de chacune des voyelles répétées 8 fois par une Française parisienne (80 réponses par voyelle)

Notons que le rapprochement effectif (acoustique ou perceptif) de la “référence” phonétique dépend dans le cas des voyelles du groupe 1 du degré de similarité qui change selon l’environnement phonétique. Ainsi, la voyelle [i], perceptivement plus éloignée du [i:] tchèque en contextes tVt et kVk, est réalisée dans ces contextes par les apprenants avec un rapprochement formantique F4-F3 moindre que chez la “référence”. De même, la voyelle [ɛ] en isolation est acoustiquement moins *similaire* au [ɛ] tchèque que le [u] ou le [i] isolé aux [u/u:] et [i:] tchèque respectivement. Ce degré de similarité moindre se traduit par une moins bonne identification par les Français natifs du [ɛ] isolé produit par les Tchèques en isolation par rapport à l’identification de leur [i] ou [u] isolés.

Ainsi au sein du groupe 1, une plus grande similarité acoustico-perceptive des voyelles du français [i, ɛ, u] avec les voyelles tchèques favorise leur maîtrise authentique que nous n’expliquons pas en termes « d’acquisition authentique » mais plutôt de « transfert positif ».

12.2.2 Groupe 2 : Voyelles *similaires* [e, o, a], divergence “symbolique”

Le deuxième groupe de voyelles acoustiquement et perceptivement *similaires* aux voyelles tchèques est celui du [e, o, a]. Ce groupe est particulier par la divergence “symbolique” selon le contexte phonétique.

12.2.2.1 Voyelle [e]

Le [e] français est dans tous les environnements *similaire*, acoustiquement et perceptivement, au [ɪ] tchèque, notamment en contextes labial et coronal. En isolation, sa similarité avec [ɪ] tchèque est moindre. En contexte palato-vélaire, il est acoustiquement *similaire* également au [ɛ/ɛ:] tchèque.

Nous avons tout d’abord voulu savoir si les Tchèques possèdent une catégorie perceptive pour le [e] isolé. Le test d’identification montre qu’elles l’ont “correctement” identifié dans 66 % des réponses seulement, avec une note moyenne de 3,9 sur 5, soit un *fit index* de 2,6. En effet, cette voyelle est majoritairement confondue avec le /ɛ/ mi-ouvert (dans 24 % des réponses, avec une note de 3,2 sur 5), comme nous pouvons le lire dans le Tableau 102. Elle est ensuite confondue avec le /œ/, dans 7,5 % des réponses et avec le /ø/, dans 2,5 % des réponses.

Stimulus	Voyelles identifiées	Pourcentage de l’identification	Echelle de confiance	<i>Fit index</i>
e	e	66	3,9	2,6
	ɛ	24	3,2	0,8
	œ	7,5	4	0,3
	ø	2,5	2	0,1

Tableau 102 : Identification par 10 Tchèques, futures enseignantes de FLE, du [e] produit en isolation par une Française parisienne, répétées 8 fois. Le stimulus est en colonne 1, les réponses en colonne 2, le pourcentage de l’identification en colonne 3, la note moyenne en colonne 4 et le *fit index*, calculé selon Guion *et al.* (2000), en colonne 5

Les données en production montrent que cette voyelle n’est pas maîtrisée. Les valeurs formantiques moyennes de la Figure 104 (page 254) montrent que le F1 du [e] final et isolé produit par les non-natives se trouve dans tous les contextes à plus de deux écarts types de la “référence”.

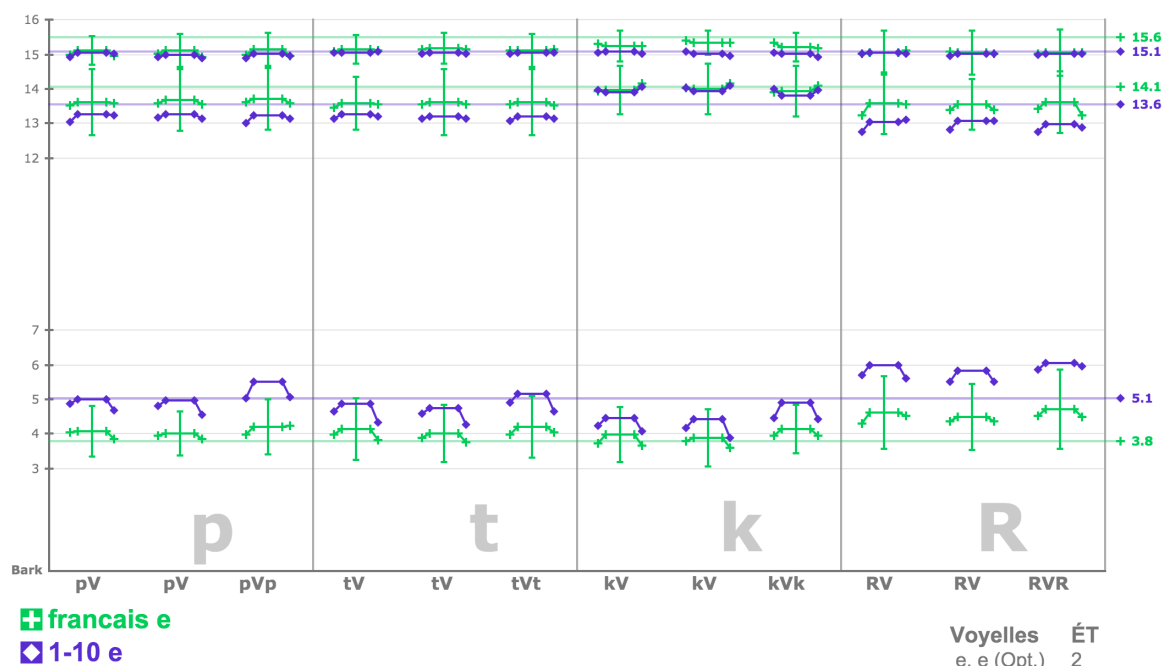


Figure 104 : Valeurs moyennes de F1, F2, F3 (en Bark) du [e] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions de dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un losange) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 2

Ainsi, malgré la forte similarité acoustique et perceptive du [e], notamment en tVt et kVk, avec la voyelle tchèque [ɪ], les futures enseignantes de FLE ne maîtrisent pas cette voyelle. Nous interprétons cet écart par la divergence “symbolique”. En regardant l’orthographe du mot *bébé*, les non-natives voient le graphème <é> qui en tchèque se prononce avec la qualité [ɛ]. Elles produisent alors un [ɛ] mi-ouvert en ne faisant pas attention à la forte similarité acoustique et perceptive que le [e] français entretient avec le [ɪ] tchèque.

L’identification du [e] isolé par les Français natifs confirme que cette voyelle n’est pas maîtrisée par les futures enseignantes de FLE : seulement dans 30 % des réponses, les stimuli sont “correctement” identifiés quoique avec une note moyenne forte (de 4,7 sur 5), soit un *fit index* de 1,4. La forte note qui juge du niveau phonétique des 30 stimuli “correctement” identifiés nous laisse penser que parmi les dix Tchèques, certaines produisent néanmoins cette voyelle avec précision. L’étude qualitative fera l’objet du chapitre 14.

12.2.2.2 Voyelle [o]

Le [o] français fait partie du deuxième groupe car il peut être *similaire*, selon le contexte, à la voyelle tchèque fermée retranscrite avec le symbole [u] (quoique les données acoustiques et perceptives ne soient pas toujours en accord sur la similarité). En effet, alors que [o] en contextes labial, coronal et palato-vélaire est perceptivement fortement *similaire* au [u/u:] tchèque, il est dans ces mêmes contextes acoustiquement *similaire* au [o/o:] tchèque. En isolation, mais dans une moindre mesure, il est perceptivement et acoustiquement *similaire* au [u/u:] tchèque.

Pour tester si les non-natives possèdent la catégorie perceptive du [o] fermé, elles ont identifié cette voyelle produite en isolation via un test perceptif décrit dans le Tableau 93 de la page 242. Le résultat montre que le [o] des tchécophones est « correctement » identifié dans seulement 56 % réponses avec une note moyenne de 3,5 sur 5, soit un *fit index* de 2. Les autres stimuli sont confondus avec le /ɔ/ ouvert et obtiennent une note moyenne de 3,5 sur 5. Notons que malgré la similarité acoustique du [o] isolé avec les [u/u:] tchèques, le [o] français n’est jamais catégorisé en tant que /u/.

Pour résumer, la catégorie perceptive pour le [o] n'est pas établie par les non-natives qui possèdent une catégorie commune pour les deux phonèmes /o/ et /ɔ/.

Pour mesurer le niveau d'authenticité dans la production de cette voyelle, nous avons comparé les formants F1 et F2 du [o] prononcé par les dix non-natives (marqué par un triangle) par rapport aux dix Françaises natives (marqué par une croix) en isolation (traits horizontaux) et en contextes [p, t, k, ʁ]. Les valeurs formantiques moyennes sont illustrées à la Figure 105 où l'écart type affiché est de un.

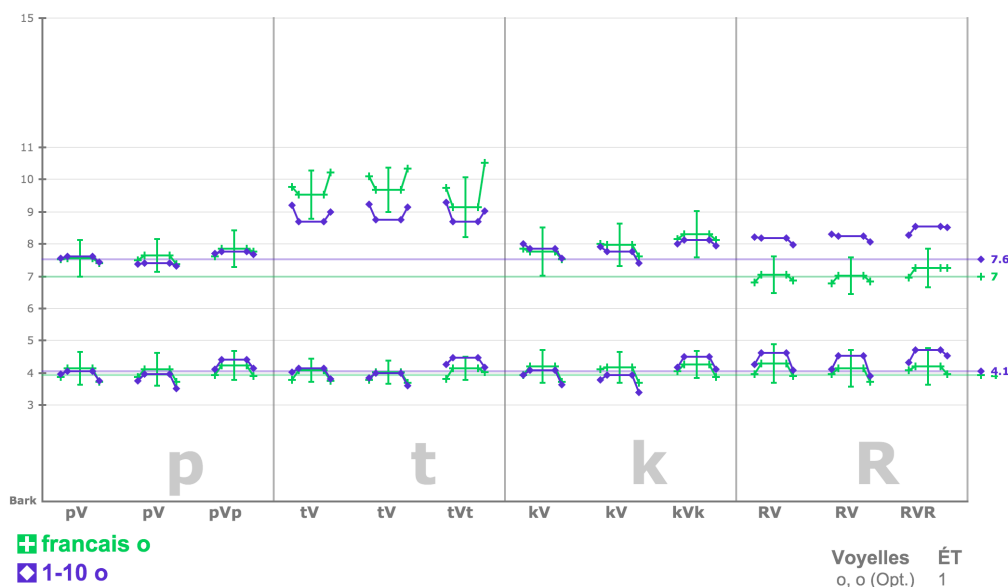


Figure 105 : Valeurs moyennes de F1, F2 du [o] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un triangle) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1

Nous remarquons à la Figure 105 que les formants F1 et F2 du [o] produit par les Tchèques en isolation et en contextes labial, coronal et palato-vélaire se trouvent à moins d'un écart type de la "référence", mis à part pour le F2 du [o] en syllabe médiane du logatome [tototot]. **La valeur de F2 plus basse par rapport à la "référence" est due à la différences de coarticulation des voyelles postérieures en contexte coronal qui provoque une plus grande élévation du F2 en français qu'en tchèque.** Cette différence n'est néanmoins pas importante car nous portons notre attention essentiellement sur la syllabe finale où [o] est maîtrisé fidèlement. En revanche, le contexte uvulaire défavorise la production fidèle car le F2 du [o] dans les trois syllabes se trouve à plus de deux écarts types de la "référence".

Malgré une réalisation par les non-natives du [o] isolé avec des formants à moins d'un écart type de la "référence", nous remarquons dans le Tableau 103 que son caractère focal est affaibli car les deux premiers formants sont plus éloignés que dans la production du [o] par des natives. En revanche, une fois les voyelles insérées en contextes [p, t, k], la distance F2-F1 des voyelles produites par les non-natives devient inférieure à celle des voyelles des natives. Ce rapprochement plus fort des deux premiers formants chez les non-natives par rapport aux locutrices natives est particulièrement saillant en contexte coronal. Cette différence s'explique par la stratégie de coarticulation du français des voyelles postérieures (avec F1 et F2 proches) qui deviennent acoustiquement centrales (F2 à peu près à mi-distance entre F1 et F3) en contexte coronal. En contexte uvulaire, cette distance est de nouveau supérieure chez les non-natives.

	0		p		t		k		R	
	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat
o	362	438	448	438	750	609	495	484	360	495

Tableau 103 : Distance acoustique F2-F1 du [o] prononcé par 10 Françaises natives (Nat) et 10 non-natives de LM tchèque (Non-nat) en isolation (0) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], 4 répétitions par locutrice

Le test d'identification du [o] isolé par les Français natifs montre que cette voyelle n'est "correctement" identifiée que dans 57 % des réponses avec une note moyenne de 4,2 sur 5, soit un *fit index* de 2,4. Dans 42 % des réponses, les stimuli sont en effet interprétés en tant que /ɔ/ mi-ouvert. Ainsi, bien que les formants F1 et F2 soient réalisés en moyenne à un écart type de la "référence", le faible rapprochement de ces deux premiers formants par rapport à la "référence" engendre une production que l'on peut considérer comme non authentique, voir non maîtrisée.

12.2.2.3 Voyelle [a]

Enfin, le [a] français est *similaire* du point de vue acoustico-perceptif à deux voyelles du tchèque. Il est soit proche des [a/a:] tchèques, maximalement en isolation, et un peu moins en contexte labial, soit aux [ɛ/ɛ:] tchèques, en contextes coronal et palato-vélaire. Cette correspondance s'explique de nouveau par la stratégie de coarticulation des voyelles postérieures du français dont le F2 s'élève de manière importante au contact des consonnes [t] ou [k], ce qui les rapproche acoustiquement des voyelles antérieures du tchèque. Ainsi malgré l'inclusion de [a] dans le groupe 2 pour indiquer une éventuelle correspondance à une voyelle tchèque retranscrite à l'aide d'un symbole différent, il serait possible d'inclure le [a] en contexte labial ou plus encore en isolation au groupe 1. Ce dernier concerne en effet les voyelles *similaires* aux voyelles tchèques selon les trois points de comparaison.

Nous avons d'abord vérifié chez les futures enseignantes de FLE l'existence de la catégorie perceptive pour [a] en isolation. Les dernières l'ont "correctement" identifié dans 97 % des réponses avec une note moyenne de 4,7, soit un *fit index* de 4,6. La catégorie perceptive du [a] isolé est alors présente dans l'esprit des futures enseignantes de FLE.

La production du [a] par les non-natives selon l'environnement phonétique est indiquée et comparée à la "référence" à la Figure 106. La figure montre l'évolution des formants F1, F2 et F3 de [a] produit par les dix Tchèques (marquée par un triangle) et les dix Françaises (marquée par une croix) en isolation (traits horizontaux) et dans des logatomes CVCVCVC où C correspond à l'une des consonnes [p, t, k, ʁ]. L'écart type affiché est de un.

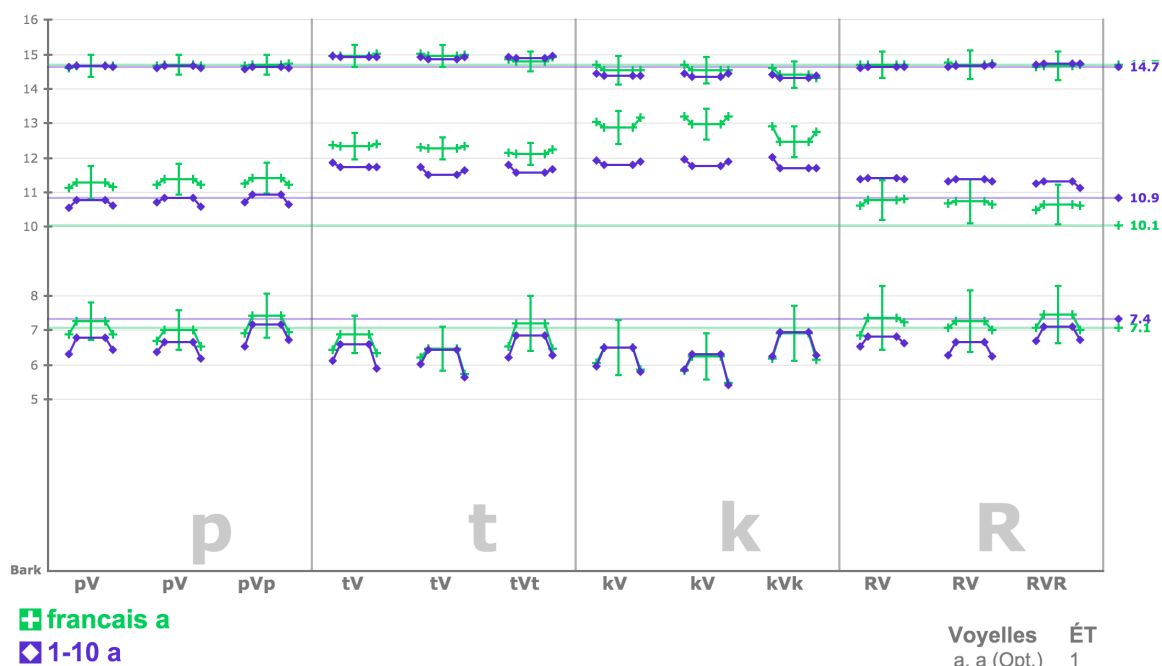


Figure 106 : Valeurs moyennes de F1, F2, F3 (en Bark) de [a] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions de dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un triangle) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1

En analysant la Figure 106, nous constatons que le premier formant du [a] produit par les non-natives se trouve à moins d'un écart type de la "référence" dans tous les contextes. En revanche, le [a] peut diverger de la "référence" par la valeur de son deuxième formant. Alors qu'en isolation et en contexte labial, le F2 se trouve à environ un écart type de la "référence", en contextes coronal, palato-vélaire et uvulaire, il se situe à environ 1,5 écart type de la "référence", mis à part en syllabe médiane des logatomes [tatatat] et [kakakak] où il est à moins de deux écarts types. Dans le cas de la production du [a], nous rencontrons alors de nouveau le phénomène de transfert, positif lorsque la voyelle se trouve en isolation et en contexte labial et négatif en contextes coronal et palato-vélaire.

La distance entre les deux premiers formants du [a] en isolation (le seul environnement où le [a] français peut être considéré comme focal) a également été vérifiée et comparée à celle observée chez les Françaises natives. Puisque le F2 du [a] des Tchèques est supérieur à celui du [a] des natives, la distance F2-F1 passe de 470 Hz en moyenne pour [a] produit par les natives du français à 631 Hz chez les non-natives, ce qui signifie que le caractère focal est atténué. Ceci n'entrave néanmoins pas son identification auprès des auditeurs français natifs qui l'identifient "correctement" dans 97 % des réponses avec une note moyenne de 4,7 sur 5 (soit un *fit index* de 4,6). Ceci signifie que le [a] isolé des non-natives est facilement identifiable. Nous constatons alors que malgré son caractère focal atténué, le [a] isolé du français est produit fidèlement par les futures enseignantes de FLE.

La voyelle moyenne [e] du groupe 2, qui peut être selon le contexte « hautement similaire » à une voyelle tchèque retranscrite avec un symbole phonétique différent [ɪ], n'est pas maîtrisée.

Le [o], perceptivement *similaire* aux [u/u:] tchèques est produit à un écart type de la "référence" dans la majorité des contextes mais il est confondu par les natifs avec /ɔ/ (du moins en isolation).

A l'inverse, la voyelle [a] est identifiée et réalisée avec précision, du moins en isolation où elle est « hautement similaire » au [a/a:] tchèque. En contextes coronal et palato-vélaire, le [a] des non-natives est réalisé avec des formants à moins de 1,5 écarts types de ceux de la "référence" mais leur identification par des auditeurs natifs n'a pas été vérifiée. Le Tableau 104 récapitule le niveau d'authenticité de [e, o, a] exprimé en production par la distance acoustique par rapport à la "référence" et le taux global d'identification (valeur du *fit index*) par les Français natifs et en perception par le taux

global d'identification par les non-natives.

Voyelle	[ɛ]					[o]					[a]				
Similarité avec une voyelle tchèque	(S)-A-P					(S)-A-P					S-A-P		(S)-A-P		
Contexte	0	p	t	k	R	0	p	t	k	R	0	p	t	k	R
Distance acoustique max. (en nb d'écarts types) des moyennes de F1, F2, (F3, F4) de la "référence"	> 2	> 2	> 2	> 2	> 2	1	1	1	1	2	1	1	1,5	1,5	<1,5
Le <i>fit index</i> de l'identification par les natifs	1,4					2,4					4,6				
Le <i>fit index</i> de l'identification par les non-natives	2,6					2					4,7				

Tableau 104 : Production et perception par 10 Tchèques des voyelles françaises [ɛ, o, a] du groupe 2 se caractérisant par une similarité acoustique (A) et perceptive (P) avec les voyelles tchèques pouvant être retranscrites selon le contexte avec un symbole différent. Le niveau en production est estimé selon la distance des formants par rapport à la "référence" (exprimée en nombre d'écarts types) et le *fit index* de l'identification par 10 Français natifs de chaque voyelle produite par 10 Tchèques (100 réponses par voyelle). Le niveau de perception est estimé par le *fit index* de l'identification par 10 Tchèques de chacune des voyelles répétées 8 fois par une Française parisienne (80 réponses par voyelle)

Ainsi, les voyelles *similaires* aux voyelles tchèques du groupe 2 sont de nouveau réalisées par un simple transfert, tout comme celles du groupe 1. Ce transfert est positif dans le cas de [a] prononcé en isolation et en contexte labial où il est fortement *similaire* aux [a/a:] tchèques. En revanche, il est négatif dès que la voyelle [a] apparaît en contextes coronal ou palato-vélaire. Dans ces contextes la voyelle est en effet nettement moins *similaire* aux [a/a:] tchèques et se rapproche davantage des [ɛ/ɛ:] tchèques.

La réalisation des voyelles moyennes du français qui sont contrairement au tchèque de deux degrés d'aperture, mi-ouvertes et mi-fermées, est influencée par la graphie qui l'emporte sur l'attention auditive. Ainsi si un Tchèque apprenant le français voit le graphème <e> qui se prononce en tchèque toujours [ɛ], ou bien <o> qui se prononce en isolation avec une qualité intermédiaire entre [o] et [ɔ] ou encore <a> qui n'est dans aucun contexte phonétique autant antériorisé que le [a] français, il ne fait alors plus attention au fait que ce son se rapproche perceptivement d'une autre voyelle présente dans le système vocalique de sa LM, à savoir le [ɪ] dans le cas du [ɛ] dans tous les contextes, le [u] dans le cas du [o] en isolation ou encore le [ɛ] dans le cas du [a] en contextes tVt ou kVk. Dans ce cas, plutôt que de procéder au transfert positif d'un son natif perceptivement (et parfois acoustiquement) *similaire* à un son maternel, il se réfère au graphème et le réalise comme il le ferait dans sa LM. De cette façon, une

divergence de la graphie et des symboles phonétiques correspondant rend généralement une perception et une production authentiques des voyelles *similaires* aux catégories maternelles impossible.

12.2.3 Groupe 3 : Voyelle [ɔ], similarité acoustique limitée

La voyelle [ɔ] est « symboliquement » *similaire* car certains auteurs tchèques utilisent ce symbole pour transcrire phonétiquement l'une des qualités vocaliques tchèques. Elle est perceptivement *similaire* dans tous les contextes au [o/o:] tchèque (avec un *fit index* de 2,5 seulement en contexte labial) mais sa similarité acoustique est limitée : en isolation, elle se trouve à moins de 1,5 écarts types du [o/o:] tchèque, en contexte labial, elle se situe à moins de deux écarts types du [a/a:] tchèque et enfin, en contextes coronal et palato-vélaire, elle est acoustiquement nouvelle. Par ce point, elle diffère des voyelles du groupe 1 et 2 qui sont dans tous les contextes acoustiquement et perceptivement *similaires* à une ou plusieurs voyelles tchèques.

Nous avons d'abord vérifié si les Tchèques possèdent la catégorie perceptive pour [ɔ] produit en isolation. Le test de perception décrit dans le Tableau 93 (page 242) montre que les non-natives ont “correctement” identifié la voyelle [ɔ] dans 55 % des réponses avec une note moyenne de 3,3 sur 5, soit un *fit index* de 1,8. Comme le fait apparaître le Tableau 105, elles l'ont essentiellement confondu avec /o/ (dans 31 % des réponses et avec une note de 3,7 sur 5), puis avec /a/ (dans 5 % des réponses, avec une note de 3,8 sur 5). Ainsi, la cible perceptive du [ɔ] isolé n'est pas acquise.

Voyelles produites par FR-NAT	Voyelles identifiées	Pourcentage de l'identification	Note	<i>Fit index</i>
ɔ	ɔ	55	3,3	1,8
	o	31	3,7	1,2
	a	5	3,8	0,2
	ã	4	2,8	0,1
	ẽ	3	2	0,1
	œ	1	3	0,03
	ø	1	3	0,03

Tableau 105 : Identification par 10 Tchèques, futures enseignantes de FLE, du [ɔ] produit en isolation par une Française parisienne, répétées 8 fois. Le stimulus est en colonne 1, les réponses en colonne 2, le pourcentage de l'identification en colonne 3, la note moyenne en colonne 4 et le *fit index*, calculé selon Guion *et al.* (2000), en colonne 5

Nous avons ensuite étudié comment les Tchèques produisent cette voyelle (peu) *similaire* à une catégorie tchèque par rapport à la “référence”. La Figure 107 illustre les valeurs formantiques moyennes du [ɔ] produit par les dix Tchèques (marqué par un triangle) et dix Françaises natives (marqué par une croix), en isolation (traits horizontaux) et en contextes [p, t, k, ʁ]. L'écart type affiché est de deux.

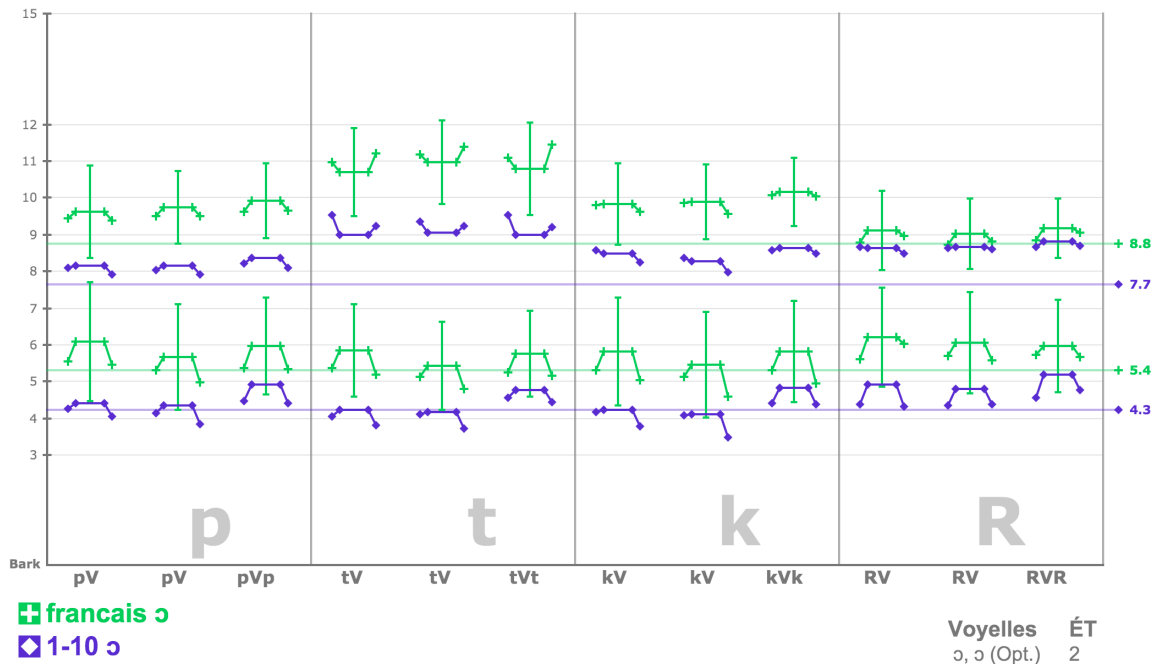


Figure 107 : Valeurs moyennes de F1 et F2 (en Bark) de [ɔ] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions de dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un triangle) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 2

Nous remarquons sur la Figure 107 que le premier formant de [ɔ] prononcé par les non-natives se trouve dans tous les contextes à plus d'un écart type de la "référence" et le deuxième formant à plus de deux écarts types, mis à part celui de [ɔ] en contexte uvulaire qui se trouve à moins d'un écart type de la "référence". Le Tableau 106 montre ensuite que le [ɔ] est réalisé avec une distance acoustique F2-F1 moindre par les non-natives que par les Françaises natives dans tous les contextes, mis à part en contexte uvulaire. Ainsi, le caractère focal du [ɔ] est renforcé en isolation et en contextes labial, coronal et palato-vélaire. Nous remarquons que le contexte coronal engendre un écartement important des formants F2/F1 du [ɔ] produit par les natives, ce qui est dû à la coarticulation des voyelles postérieures avec la consonne [t] qui élève fortement le deuxième formant de ces dernières.

Il est intéressant de noter que même en isolation où la distance entre le [ɔ] français et le [o] tchèque est inférieur à deux écarts types (c'est alors le contexte où les deux sons sont acoustiquement maximale*ment similaires*), les Tchèques produisent le [ɔ] français avec un F1 et F2 si bas qu'il se situe à plus de deux écarts types de la "référence". Dans ce cas, il ne s'agit alors pas d'un transfert de la LM vers la LE mais plutôt du respect de la loi de position qui exige la fermeture de la voyelle isolée. Ainsi le [ɔ] mi-ouvert (plus proche du [ɔ] isolé tchèque ayant une grande aperture) est réalisé comme un [o] mi-fermé (plus éloigné du [o] isolé tchèque).

F2-F1	0		p		t		k		R	
	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat
ɔ	475	430	588	485	835	635	651	544	448	517

Tableau 106 : Distance acoustique F2-F1 (en Hz) du [ɔ] prononcé par 10 Françaises natives (Nat) et 10 Non-natives (Non-nat) en isolation (0) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], 4 répétitions par locutrice

L'identification du [ɔ] isolé par les auditeurs natifs a montré que dans seulement 57 % des réponses, les stimuli avaient été "correctement" identifiés avec une note moyenne de 4,1, soit un *fit*

index de 2,3. Dans 41 % des réponses, les stimuli avaient été en revanche identifiés en termes de /o/ mi-fermé. Ainsi, cette voyelle n'est pas maîtrisée en production par les Tchèques.

Pour la production de la voyelle [ɔ], les locutrices non-natives s'écartent d'une manière importante de la "référence". En contexte labial, où le [ɔ] est acoustiquement *similaire* aux [a/a:] tchèques et en contextes coronal et palato-vélaire, où la voyelle [ɔ] est acoustiquement *nouvelle*, les futures enseignantes produisent cette voyelle avec le formant F2 qui se situe à plus de deux écarts types de la "référence" et un caractère focal fort. En isolation, où [ɔ] est acoustiquement *similaire* au [o/o:] tchèque, les Tchèques visent une plus grande fermeture de la voyelle et elle est perceptivement confondue par les Français natifs avec la voyelle mi-fermée /o/. Ce résultat peut s'expliquer par la connaissance et l'application des règles orthoépiques du français standard qui interdisent la distribution du /o/ en syllabe finale CV. Ainsi, il ne doit pas être interprété comme de l'échec dans l'apprentissage du FLE.

Le Tableau 107 récapitule le niveau d'authenticité du [ɔ] exprimé en production par la distance acoustique par rapport à la "référence" et le taux global d'identification (valeur du *fit index*) par les Français natifs et en perception par le taux global d'identification chez les non-natives.

Voyelle	[ɔ]				
Similarité avec une voyelle tchèque	S-A-P	(S)-A-P	S-P		
Contexte	0	p	t	k	R
Distance acoustique max. (en nb d'écarts types) des moyennes de F1, F2, (F3, F4) de la "référence"	> 2	> 2	> 2	> 2	<1,5
Le <i>fit index</i> de l'identification par les natifs	2,3				
Le <i>fit index</i> de l'identification par les non-natifs	1,8				

Tableau 107 : Production et perception par 10 Tchèques du [ɔ] se caractérisant par une similarité perceptive (P) et acoustique restreinte (A) avec les voyelles tchèques pouvant être retranscrites selon le contexte avec un symbole différent. Le niveau en production est estimé selon la distance des formants par rapport à la "référence" (exprimée en nombre d'écarts types) et le *fit index* de l'identification par 10 Français natifs de chaque voyelle produite par 10 Tchèques (100 réponses par voyelle). Le niveau de perception est estimé par le *fit index* de l'identification par 10 Tchèques de chacune des voyelles répétées 8 fois par une Française parisienne (80 réponses par voyelle)

Ainsi le [ɔ], indépendamment du degré et type de similarité qu'il entretient avec la/les voyelles tchèques selon le contexte, n'est pas maîtrisé par les futures enseignantes de FLE de LM tchèque.

12.2.4 Groupe 4 : Voyelle [œ], "symboliquement" *nouvelle*, similarité acoustique limitée

La voyelle [œ] a un statut particulier : elle est retranscrite avec un symbole de l'API nouveau, inutilisé par le système vocalique tchèque mais elle est perceptivement *similaire* dans tous les contextes à une seule catégorie : les [ɛ/ɛ:] tchèques. Elle est également acoustiquement proche de cette

voyelle en contextes labial, palato-vélaire et maximale en contexte coronal où les formants F1 et F2 du [œ] se trouvent à moins d'un écart type du [ɛ/ɛ:] tchèque.

La perception du [œ] isolé par les Tchèques a été vérifiée au moyen du test d'identification décrit dans le Tableau 93 de la page 242: dans 74 % des réponses, les stimuli ont été “correctement” catégorisés avec une note moyenne de 3,6 sur 5, soit un *fit index* de 2,7. Les autres stimuli ont été confondus, comme nous pouvons le voir dans le Tableau 108, avec /ø/ dans 13 % des réponses, avec /ɛ/ dans 11 % des réponses et enfin avec /e/ dans 2 % des réponses. Ainsi, les Tchèques ne catégorisent pas cette voyelle en isolation comme les natifs du français. En effet, le résultat du test d'identification en 11.1 montre qu'un natif a identifié les voyelles françaises isolées à 100 %.

Voyelles produites par FR-NAT	Voyelles identifiées	Pourcentage de l'identification	Echelle de confiance	<i>Fit index</i>
œ	œ	74	3,6	2,7
	ø	13	3,4	0,4
	ɛ	11	2,9	0,3
	e	2	1,5	0,03

Tableau 108 : Identification par 10 Tchèques, futures enseignantes de FLE, du [œ] produit en isolation par une Française parisienne, répétées 8 fois. Le stimulus est en colonne 1, les réponses en colonne 2, le pourcentage de l'identification en colonne 3, la note moyenne en colonne 4 et le *fit index*, calculé selon Guion *et al.* (2000), en colonne 5

La Figure 108 illustre comment cette voyelle est réalisée (en termes des formants moyens) par les non-natives (marqué par un triangle), en comparaison avec la “référence” (marqué par une croix) en isolation et dans des logatomes CVCVCVC où C correspond à la consonne [p, t, k, ʀ].

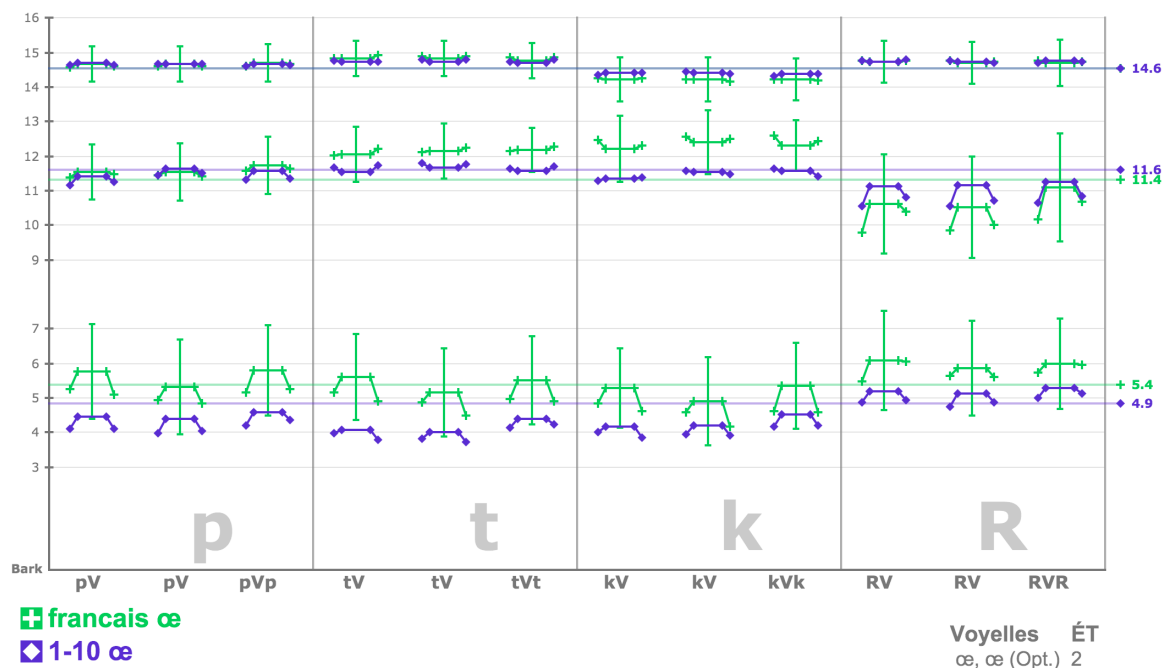


Figure 108 : Valeurs moyennes de F1, F2 et F3 du [œ] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʀ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un triangle) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 2

[illegible]

Malgré une réalisation des formants du [œ] isolé par les Tchèques à moins d'un écart type de ceux du [œ] isolé des natifs, les auditeurs français l'identifient en termes de /œ/ dans 55 % des réponses, avec une note moyenne de 4,5 sur 5 (soit un *fit index* de 2,5) et en termes de /ø/ dans 43 % des réponses, avec une note moyenne de 3.7 sur 5, soit un *fit index* de 1.6.

Le niveau de maîtrise en production du [œ] (mesuré selon la distance acoustique des formants par rapport à la “référence”) est inversement proportionnel au degré de similarité que [œ] entretient avec le [ɛ/ɛ:] tchèque. Alors que l’approximation de la cible acoustique est maximale pour [œ] isolé (comme indiqué dans le Tableau 109) qui n’est que perceptivement *similaire* au [ɛ/ɛ:] tchèque, elle est minimale pour [œ] en contextes coronal et palato-vélaire où la similarité est non seulement perceptive mais également acoustique. [œ] en contexte labial entretient une similarité perceptive moindre avec [ɛ/ɛ:] tchèque, ce qui se traduit par une plus grande approximation de la “référence” en termes des valeurs formantiques uniquement dans le cas où [œ] se trouve en contextes coronal ou palato-vélaire.

Voyelle	[œ]				
Similarité avec une voyelle tchèque	P	A-P			
Contexte	0	p	t	k	R
Distance acoustique max. (en nb d'écarts types) des moyennes de F1, F2, (F3, F4) de la "référence"	1	< 2	2	2	1
Le <i>fit index</i> de l'identification par les natifs	2,5				
Le <i>fit index</i> de l'identification par les non-natifs	2,7				

Tableau 109 : Production et perception par 10 Tchèques du [œ] se caractérisant par un symbole *nouveau* et une similarité perceptive (P) et acoustique (A) limitée aux contextes labial et coronal. Le niveau en production est estimé selon la distance des formants par rapport à la “référence” (exprimée en nombre d’écarts types) et le *fit index* de l’identification par 10 Français natifs de chaque voyelle produite par 10 Tchèques (100 réponses par voyelle). Le niveau de perception est estimé par le *fit index* de l’identification par 10 Tchèques de chacune des voyelles répétées 8 fois par une Française parisienne (80 réponses par voyelle)

Les résultats montrent que la voyelle [œ] n'est généralement pas maîtrisée en isolation. Ce phénomène peut de nouveau s'expliquer par la loi de position et par conséquent ne doit pas être considéré comme de l'échec dans l'apprentissage phonétique du FLE.

12.2.5 Groupe 5 : Voyelles [y, ø], “symboliquement” et acoustiquement *nouvelles*, similarité perceptive limitée

Le Tableau 110 montre dans quels environnements les voyelles [y, ø] peuvent être considérées comme *similaires* aux voyelles tchèques (cases grisées). Notons alors que la similarité n’est que perceptive, contrairement à la voyelle [œ] du groupe 4. Le son tchèque auquel la voyelle française renvoie est indiqué pour chaque contexte séparément.

Voyelle/contexte	0	p	t	k
y	i	i:		i
ø		ε/ε:	u/u:	

Tableau 110 : Les voyelles françaises *nouvelles* [y, ø] selon le contexte (nul (0) ou [p,t,k]) du point de vue “symbolique” (S), acoustique (A) et perceptif (P). Les voyelles tchèques perceptivement *similaires* sont entre parenthèses. Le contexte qui rend la voyelle *nouvelle* selon les trois critères est grisé

Le Tableau 110 montre que :

- la voyelle [y] est perceptivement *similaire* au [ɪ] bref tchèque en contexte nul et palato-vélaire et au [i:] long tchèque en contexte labial
- la voyelle [ø] est perceptivement *similaire* au [ε/ε:] tchèque en contexte labial et au [u/u:] tchèque en contexte coronal
- dans les autres contextes, ces voyelles sont *nouvelles*

Afin de savoir si les Tchèques ont établi une catégorie perceptive pour [y] isolé, nous leur avons proposé un test d’identification décrit dans le Tableau 93. Le résultat montre qu’elles l’identifient “correctement” dans 98 % des réponses avec une note moyenne de 4,4, soit un *fit index* de 4,3. Ainsi, la maîtrise de la catégorie perceptive [y] isolé par les futures enseignantes de FLE est authentique. Notons que l’établissement de la catégorie perceptive pour [ø] en contextes labial et coronal (seuls contextes où [ø] est une voyelle *similaire* à une catégorie maternelle) n’a pas été vérifié.

Pour juger de l’authenticité de la production de ces voyelles dans les contextes où elles sont perceptivement *similaires* aux voyelles tchèques, nous avons comparé les valeurs de leurs trois premiers formants par rapport à la « référence ». Le [y] est comparé à la Figure 109 (page 265) et le [ø] à la Figure 110 (page 266). Les valeurs formantiques moyennes y sont calculées à partir de 4 répétitions des voyelles isolées (traits horizontaux) et en contextes [p, t, k, ʁ] (4 répétitions*10 locutrices). L’écart type affiché est de un.

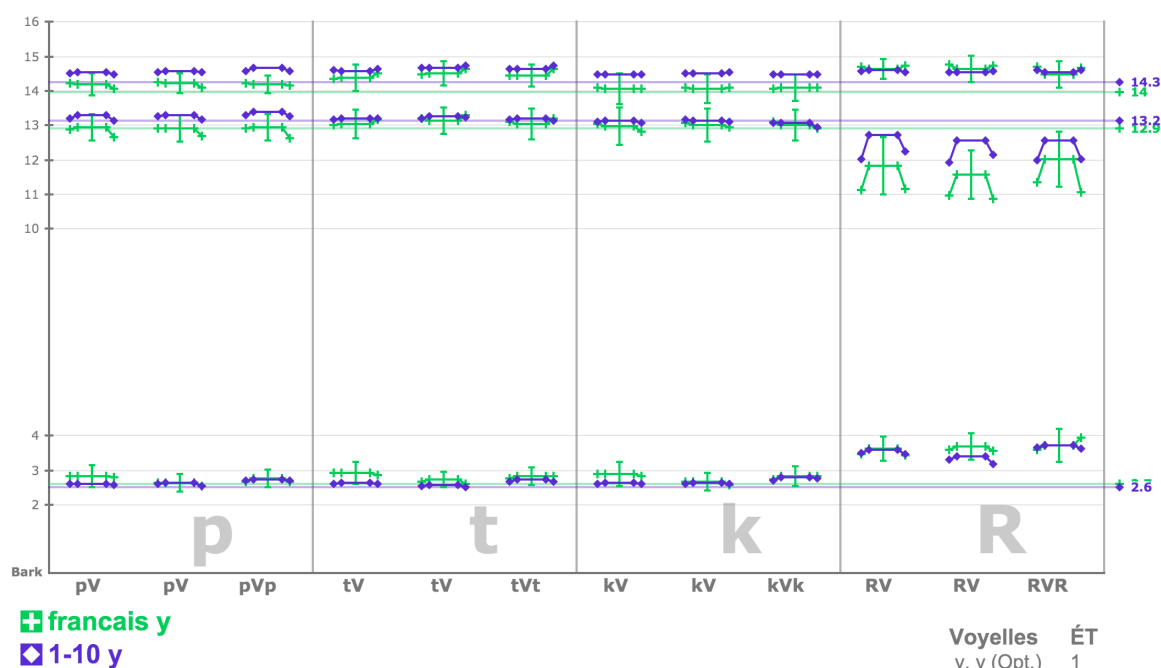


Figure 109 : Valeurs moyennes de F1, F2 et F3 du [y] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un triangle) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1

Nous remarquons sur la Figure 109 que les formants F1, F2, F3 du [y] produit par les non-natives se trouvent à un écart type ou moins de la “référence” dans tous les contextes, mis à part en contexte labial où le F3 se situe à environ deux écarts types. Ce résultat n'est que partiellement en accord avec les prédictions basées sur le SLM qui ne prévoit pas l'authenticité pour les sons *similaires*.

En ce qui concerne le caractère focal du [y] marqué par le rapprochement du deuxième et troisième formant, le Tableau 111 montre qu'il est plus faible chez les non-natives que chez les natives, mis à part en contexte uvulaire.

F3-F2	0		p		t		k		R	
	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat	Nat	Non-nat
y	417	470	507	561	587	628	426	585	1039	784

Tableau 111 : Distance acoustique F3-F2 du [y] prononcé par 10 Françaises natives (Nat) et 10 non-natives (Non-nat) en isolation (0) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], 4 répétitions par locutrice

Pour savoir si l'écartement plus important de F2 et F3 du [y] isolé produit par les non-natives dégrade sa perception auprès des auditeurs français natifs, nous avons effectué un test d'identification par ces derniers. Le test est décrit dans le Tableau 93 à la page 242. Le résultat montre que [y] isolé est « correctement » identifié dans 96 % des réponses avec une note moyenne de 4,6 sur 5, soit un *fit index* de 4,4. Ainsi, nous pouvons considérer que le [y] isolé est produit authentiquement par les tchécophones.

La Figure 110 illustre ensuite la production de la voyelle [ø]. Nous ne nous intéressons qu'aux contextes labial et coronal où [ø] est une voyelle perceptivement *similaire* à une catégorie maternelle. Nous remarquons que dans ces contextes, les formants F1, F2, F3 se réalisent à un écart type ou moins de la “référence”, ce qui atteste d'une production a priori authentique. La perception par les auditeurs

français natifs n'est néanmoins pas vérifiée car seules les voyelles isolées ont fait l'objet du test d'identification.

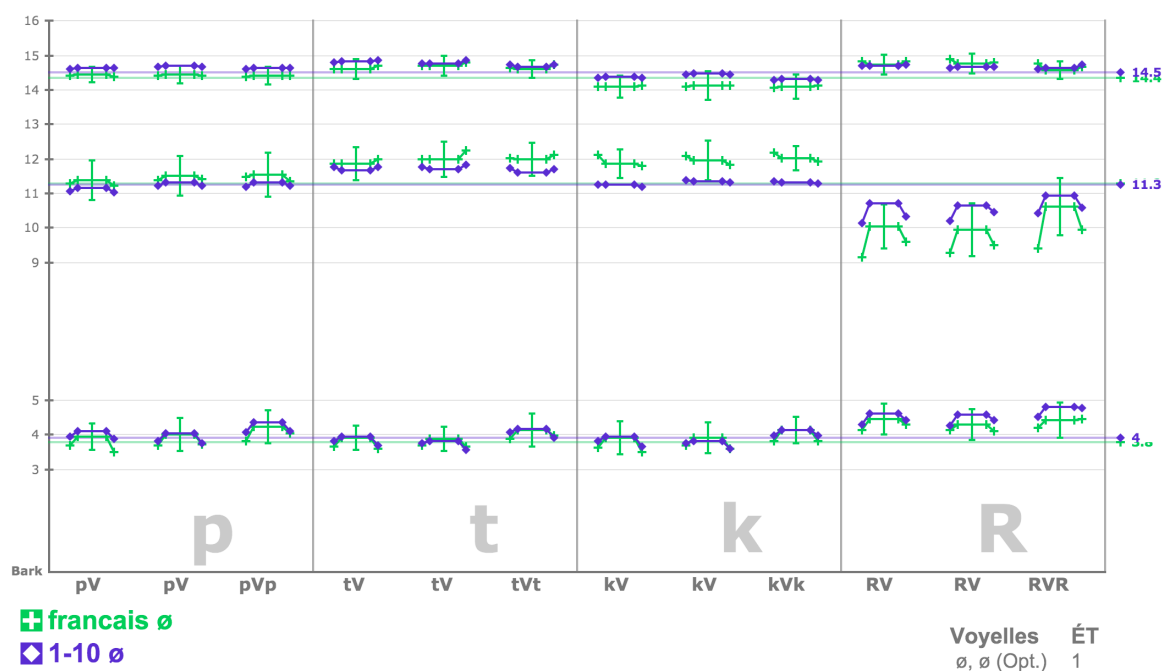


Figure 110 : Valeurs moyennes de F1, F2 et F3 du [ø] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un triangle) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1

L'approximation de la cible de la voyelle [y], acoustiquement et « symboliquement » nouvelle, est inversement proportionnelle au degré de sa similarité perceptive avec les voyelles tchèques. Alors que [y] en contexte labial est perceptivement maximale *similaire* (au [i:] tchèque), il est produit moins fidèlement dans ce contexte que dans les autres contextes où il entretient une similarité perceptive moindre (avec le [ɪ] tchèque). Contrairement aux attentes, la cible de la voyelle [ø] en contexte palato-vélaire où cette dernière est perceptivement *nouvelle* est moins atteinte qu'en contextes labial et dental où [ø] est perceptivement *similaire* aux voyelles tchèques.

Le Tableau 112 montre le niveau d'authenticité de l'acquisition de ces voyelles selon le contexte phonétique. La maîtrise en production est exprimée par la distance acoustique maximale entre les formants des voyelles des non-natives et la "référence" (en nombre d'écarts types) et par l'identification par les auditeurs français natifs (par la valeur du *fit index*). La maîtrise en perception est exprimée par la valeur du *fit index* qui renseigne sur l'identification des voyelles produites par les locutrices françaises natives et perçues par les non-natives.

Voyelle	[y]					[ø]				
Similarité avec une voyelle tchèque	P		/	P		/	P		/	
Contexte	0	p	t	k	R	0	p	t	k	R
Distance acoustique max. (en nb d'écarts types) des moyennes de F1, F2, (F3, F4) de la "référence"	1	2	1	1	1	1	1	1	2	1
Le <i>fit index</i> de l'identification par les natifs	4,4					3,7				
Le <i>fit index</i> de l'identification par les non-natives	4,3					3,4				

Tableau 112 : Production et perception par 10 Tchèques des [y, ø] se caractérisant par un symbole *nouveau* et, selon le contexte, une similarité perceptive (P). Le niveau en production est estimé selon la distance des formants par rapport à la "référence" (exprimée en nombre d'écarts types) et le *fit index* de l'identification par 10 Français natifs de chaque voyelle produite par 10 Tchèques (100 réponses par voyelle). Le niveau de perception est estimé par le *fit index* de l'identification par 10 Tchèques de chacune des voyelles répétées 8 fois par une Française parisienne (80 réponses par voyelle). Les contextes où ces voyelles sont *similaires* sont grisés

12.3 Voyelles orales du français en contexte uvulaire (analyse formantique)

La similarité des voyelles du français et du tchèque en contexte uvulaire n'a pas pu être testée car /ʁ/ ne fait pas partie du système consonantique du tchèque. Ainsi, en s'appuyant sur le SLM, nous

n'avons pas pu faire des prédictions à propos de la possibilité de la maîtrise authentique par les tchécophones des voyelles orales du français dans ce contexte.

La Figure 111 illustre les triangles vocaliques F1/F2 et F2/F3 (en Bark) des dix voyelles orales du français prononcées en contexte uvulaire par dix Tchèques (marquées par un triangle) et dix Françaises natives (marquées par une croix) qui servent de “référence”. Les ellipses de dispersion sont tracées à un écart type.



Figure 111 : Triangles vocaliques F1/F2, F2/F3 (en Bark) des voyelles orales du français produites en syllabe finale des logatomes RVRVRVR par dix Françaises (marquées par un rond) et dix Tchèques (marquées par un triangle), 4 répétitions. Les ellipses de dispersion sont tracées à un écart type

La Figure 111 montre que les locutrices non-natives réalisent les voyelles postérieures et antérieures labiales en contexte uvulaire avec un F2 plus élevé par rapport à la “référence”. Les voyelles antérieures étirées [i] et [ε] et les voyelles labiales [y], [ø] et [u] des Tchèques se trouvent à environ un écart type de la “référence”. En revanche, les voyelles [œ], [ɔ] et [a] sont à plus d’un écart type et les voyelles [e] et [o] sont à plus de deux écarts types de la “référence” et donc a priori non maîtrisées. Notons que la voyelle finale [e] n’est maîtrisée dans aucun environnement phonétique car

le F1 se trouve systématiquement à plus de deux écarts types de la “référence”. En revanche, la voyelle [o] est maîtrisée dans tous les contextes, avec les formants F1 et F2 à moins de deux écarts types de la “référence”, mis à part en contexte uvulaire. A l’inverse, le contexte uvulaire favorise la réalisation du [ɔ] par les Tchèques. Les formants F1 et F2 s’y trouvent en effet à moins de deux écarts types de la “référence” alors que dans les autres contextes, [ɔ] mi-ouvert n’est à priori pas maîtrisé.

12.4 Voyelles orales du français non maîtrisées par les futures enseignantes tchèques

L’étude du chapitre 12 montre que certaines voyelles orales étaient sont par les futures enseignantes de FLE de LM tchèque selon le contexte avec l’un de leurs formants de la syllabe finale ou de la voyelle isolée à deux écarts types ou plus de la “référence”. Cela concerne les syllabes en position finale ou les voyelles isolées et varie selon le contexte. Ces voyelles se rapprochent ainsi d’une autre catégorie française. Les voyelles qui ne sont pas maîtrisées par les futures enseignantes tchèques sont récapitulées dans le Tableau 113.

Qualité vocalique visée	Produite par les enseignantes avec le timbre :	En contexte :
[e]	[ɛ]	- isolation - syllabe finale pVp, tVt, kVk, RVR ²⁰
[œ]	[ø]	- syllabe finale pVp, tVt ²¹ , kVk
[ɔ]	[o]	- isolation ²² - syllabe finale pVp, tVt, kVk
[o]	[ɔ]	- syllabe finale RVR ²³

Tableau 113 : Voyelles du français non maîtrisées par les futures enseignantes de FLE

²⁰ En suivant les règles orthopéiques du français standard sur la distribution des timbres des voyelles moyennes en fonction de la **structure syllabique**, ce résultat ne doit pas être considéré comme de l’échec

²¹ En suivant les règles orthopéiques du français standard sur la distribution des timbres des voyelles moyennes en fonction du **contexte phonémique**, ce résultat ne doit pas être considéré comme de l’échec

²² En suivant les règles orthopéiques du français standard sur la distribution des timbres des voyelles moyennes en fonction de la **structure syllabique**, ce résultat ne doit pas être considéré comme de l’échec

²³ En suivant les règles orthopéiques du français standard sur la distribution des timbres des voyelles moyennes en fonction du **contexte phonémique**, ce résultat ne doit pas être considéré comme de l’échec

Le Tableau 113 montre que :

- Le [e] des tchécophones est dans tous les contextes acoustiquement proche du [ɛ] des Françaises natives,
- le [ɔ] en isolation, en contextes labial, coronal et palato-vélaire est acoustiquement proche du [o] des natives,
- le [o] est acoustiquement proche du [ɔ] des natives en contexte uvulaire,
- le [œ] en contextes labial, coronal et palato-vélaire est acoustiquement proche du [ø] des natives.

Ainsi, la difficulté maximale dans l'apprentissage des voyelles du français par les Tchèques réside dans les contrastes entre les voyelles moyennes mi-ouvertes et mi-fermées, c'est-à-dire entre o/ɔ, e/ɛ et ø/œ. La paire e/ɛ dont le contraste est fragilisé dans tous les contextes, est constituée de deux voyelles *similaires* aux voyelles du tchèque : [e] étant *similaire* selon le critère perceptif et acoustique mais avec une divergence “symbolique” au [ɪ] et [ɛ] étant « hautement similaire » selon les trois critères aux [ɛ/ɛ:]. La paire e/ɛ est réalisée au profit de la voyelle « hautement similaire » [ɛ]. La paire o/ɔ dont le contraste est fragilisé dans tous les contextes, est constituée de deux voyelles *similaires* : [o] *similaire* perceptivement et acoustiquement, [ɔ] *similaire* perceptivement mais avec une similarité acoustique limitée. Les deux voyelles peuvent correspondre à des voyelles tchèques retranscrites avec un symbole phonétique différent ([u, a] respectivement). La paire o/ɔ est ainsi réalisée au profit de la variante mi-fermée [o] en isolation et en contextes labial, coronal et palato-vélaire. En contexte uvulaire, la paire se réalise au profit du [ɔ]. Enfin, le contraste de la paire des voyelles ø/œ est fragilisé en contextes labial, coronal et palato-vélaire où la paire se réalise au profit de la voyelle [ø] qui est acoustiquement nouvelle. Même si les deux voyelles sont retranscrites avec un symbole phonétique nouveau, [œ] entretient une similarité perceptive et acoustique avec le [ɛ/ɛ:] tchèque.

Le Tableau 114 récapitule la façon dont les voyelles isolées produites par les Tchèques sont perçues par les Français natifs. Les voyelles y sont classées selon la valeur du *fit index*, des mieux identifiées par les natifs aux difficilement identifiables qui correspondent en effet aux voyelles moyennes.

Voyelles produites par T1-T10	Voyelles les plus souvent identifiées	Pourcentage de l'identification	Echelle de confiance	Fit index
i	i	100	4,9	4,9
u	u	99	4,9	4,9
a	a	97	4,7	4,6
y	y	96	4,6	4,4
ɛ	ɛ	92	4,5	4,1
ø	ø	78	4,7	3,7
	œ	22	4,8	1
œ	œ	55	4,5	2,5
	ø	43	3,7	1,6
o	o	57	4,2	2,4
	ɔ	42	3,9	1,6
ɔ	ɔ	57	4,1	2,3
	o	41	4,4	1,8
e	e	30	4,7	1,4
	ɛ	68	4,5	3,1

Tableau 114 : Identification par 10 Français natifs des 10 voyelles orales produites en isolation par 10 futures enseignantes de FLE. Les stimuli sont en colonne 1, les réponses en colonne 2, le pourcentage de l'identification en colonne 3, la note moyenne en colonne 4 et le *fit index*, calculé selon Guion *et al.* (2000), en colonne 5

Le résultat du test d'identification du Tableau 114 révèle que le contraste entre les voyelles moyennes du français n'est généralement pas maîtrisé par les Tchèques. Même si la voyelle [ɛ] « hautement similaire » aux [ɛ/ɛ:] tchèques et la voyelle *nouvelle* [ø] sont identifiées par les natifs en termes de la catégorie visée avec plus de facilité, le [e] et le [œ] sont en partie interprétés par respectivement les /ɛ/ et /ø/ et donc les contrastes e/ɛ et ø/œ sont fragilisés. De même, le contraste entre o/ɔ n'est pas maîtrisé.

Le Tableau 115 de la page 272 récapitule enfin l'identification des dix voyelles orales du français, produites par une Française native et perçues par les dix Tchèques. Elles sont classées de nouveau selon la valeur du *fit index* qui englobe le pourcentage de l'identification et la note obtenue. Nous remarquons que les voyelles moyennes sont moins bien identifiées que les autres voyelles. Il faut noter également qu'au sein des paires des voyelles moyennes isolées, le timbre mi-fermé est plus facilement identifié que le timbre mi-ouvert. L'identification de la qualité mi-fermée plutôt que mi-ouverte des voyelles moyennes en isolation est probablement guidée par la loi de position et ne doit pas être considéré comme de l'échec.

Voyelles produites par FR-NAT	Voyelles les plus souvent identifiées	Proportion de l'identification	Echelle de confiance	Fit index
i	i	99	4,8	4,7
a	a	100	4,7	4,7
u	u	96	4,6	4,4
y	y	98	4,4	4,3
ø	ø	84	4,1	3,4
	œ	12,5	3,3	0,4
œ	œ	74	3,6	2,7
	ø	12,5	3,4	0,4
e	e	66	3,9	2,6
	ɛ	24	3,1	0,7
ɛ	ɛ	61	3,8	2,3
	e	28	3,7	1
o	o	56	3,5	2
	ɔ	44	3,4	1,5
ɔ	ɔ	55	3,3	1,8
	o	31	3,7	1,2

Tableau 115 : Identification par 10 Tchèques, futures enseignantes de FLE, des 10 voyelles orales produites en isolation par une Française parisienne, répétées 8 fois. Les stimuli sont en colonne 1, les réponses en colonne 2, le pourcentage de l'identification en colonne 3, la note moyenne en colonne 4 et le *fit index*, calculé selon Guion *et al.* (2000), en colonne 5

Le Tableau 115 montre que les locutrices non-natives ne perçoivent pas le contraste entre les voyelles moyennes o/ɔ et que la perception du contraste e/ɛ est également fragilisée même si dans plus de 60 % des réponses les stimuli sont “correctement” identifiés. La catégorie perceptive la plus fortement établie parmi les voyelles moyennes du français est le /ø/. Ce résultat correspond aux attentes du SLM qui prévoit une facilité dans l'acquisition des voyelles *nouvelles*.

Le résultat à propos de l'identification des voyelles orales françaises par des tchécophones peut être comparé avec l'étude analogue de (Dohalska-Zichova, 1980). L'auteur a effectué un test d'identification des voyelles tchèques et françaises. Au total 250 voyelles produites en isolation (10 voyelles tchèques*5 locuteurs tchèques natifs*2 répétitions et 15 voyelles françaises* 5 locuteurs français natifs* 2 répétitions) ont été identifiées par 5 auditeurs français natifs (parlant également le tchèque) et 5 auditeurs tchèques natifs (parlant également le français). La consigne était la suivante : « Indiquez toujours la voyelle tchèque ou française qui vous semble être la plus proche du son produit. » Les auditeurs tchèques montrent une facilité à identifier « correctement » les voyelles

françaises fermées [i, y, u], ce qui est en accord avec nos résultats. En revanche, la voyelle [a] est contrairement à nos résultats moins souvent interprétée en termes de la voyelle française antérieure car elle est largement confondue avec le /a/. Cette différence des résultats s'explique par notre choix de ne pas tester l'identification du /a/ postérieur qui n'a plus de statut phonologique en français actuel. En revanche, les résultats se rejoignent de nouveau dans la difficulté à « correctement » identifier les voyelles moyennes mi-ouvertes qui sont largement confondues avec leurs homologues mi-fermées. Les pourcentages exacts d'identifications « correctes » diffèrent de nos résultats ce qui peut s'expliquer par l'identification des voyelles françaises au travers des seules catégories françaises dans notre test et non pas au travers des voyelles des deux langues, comme chez Dohalska-Zichova.

12.5 Conclusion partielle

La perception (examinée au travers des tests d'identification) et production (étudiée sur le plan acoustique) des voyelles orales du français par des futures enseignantes tchèques peut-être décrite essentiellement en termes de deux processus :

1. Les voyelles acoustiquement et perceptivement *similaires* aux voyelles tchèques, retranscrites avec un symbole phonétique exploité par le système phonético-phonologique du tchèque sont le plus souvent produites par la catégorie de la LM la plus proche. Si la voyelle du français est « hautement similaire » à la voyelle du tchèque, elle peut alors être produite et perçue authentiquement par les tchécophones. Si en revanche une plus grande distance acoustico-perceptive sépare la voyelle du français de la voyelle du tchèque correspondante, sa réalisation n'est généralement plus authentique. Ce phénomène correspond aux limites de l'apprentissage expliquées par Flege (1987d) en termes de *mécanisme de classification par équivalence*. Ainsi, pour ce type de voyelles, une plus grande similarité est en général favorable à l'atteinte de la cible en L2. Cependant, il est important de souligner l'immense influence de la graphie qui survient lors d'un apprentissage formel, décrit chez les apprenants du français de LM tchèque par exemple par Fenclova (2003). L'auteure prend l'exemple de la voyelle [e] du français qui détient, comme le montrent nos données, une grande similarité acoustico-perceptive avec le [ɪ] bref tchèque. Lorsque les Français épellent l'alphabet, les apprenants tchèques débutants qui découvrent le système sonore du français (sans en connaître l'écriture) entendent [a], [bɪ], [sɪ], [dɪ], etc. En revanche dès lors qu'ils prennent connaissance du symbole phonétique [e] et de l'écriture de ce son, qui est noté souvent par le graphème <e>, ils oublient sa similarité acoustico-perceptive avec le [ɪ] tchèque et le perçoivent et le réalisent au travers de la qualité vocalique [ɛ] à laquelle renvoie systématiquement le graphème <e> en tchèque. De cette façon, lorsque la graphie est divergente, le transfert des caractéristiques acoustico-perceptives d'un son *similaire* à une catégorie maternelle n'a généralement pas lieu car cette similarité n'est plus perçue.
2. L'apprentissage phonétique des voyelles acoustiquement *nouvelles*, toujours retranscrites avec un symbole phonétique inutilisé en tchèque, est possible. Contrairement aux voyelles *similaires* à un son de la LM selon les trois critères de comparaison, une plus grande similarité (toujours uniquement perceptive pour ce type de sons) peut freiner l'approximation de la "référence". Nos résultats ne soutiennent cependant que partiellement le SLM car les voyelles *nouvelles* ne sont pas systématiquement perçues et produites avec authenticité par les futures enseignantes.

Ces résultats sont ainsi en partie en accord avec les prédictions basées sur le modèle d'apprentissage phonétique SLM.

Le chapitre 12 a ensuite révélé que l'approximation de la cible des voyelles françaises produites par les Tchèques varie d'un contexte à l'autre (car la similarité interlangue dépendait du contexte) et que les contrastes entre les voyelles moyennes du français ne sont généralement pas maîtrisés. Le triangle F1/F2 (en Bark) de la Figure 112 illustre la grande proximité acoustique des voyelles moyennes du français produites par les futures enseignantes de FLE en isolation (triangle

maximal, marqué en orange par un rond), en contextes labial (marqué en bleu par un triangle), coronal (marqué en rouge par une croix), palato-vélaire (marqué en violet par un losange) et uvulaire (marqué en vert par un triangle vers le bas). Notons que seules les voyelles [œ] et [ø] en isolation sont produites avec une différence acoustique qui soit susceptible d'être perçue.

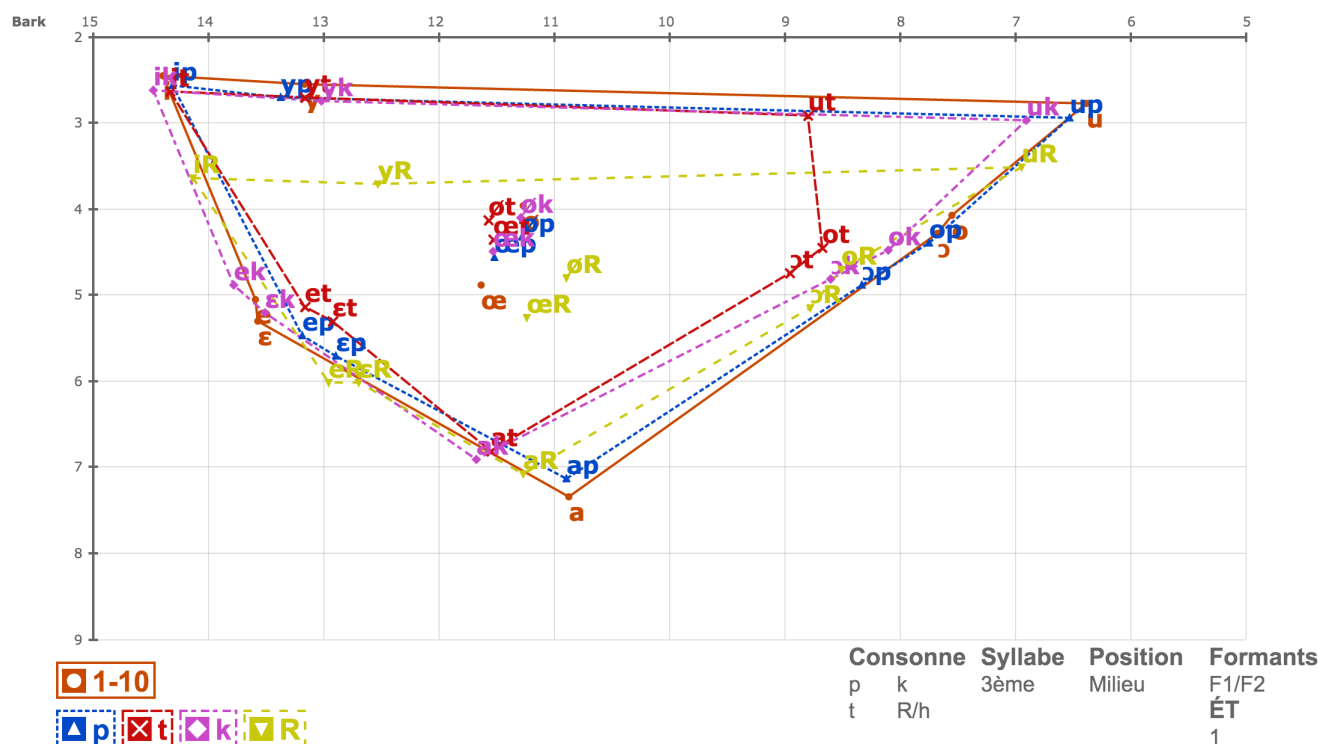


Figure 112 : Triangle F1/F2 (en Bark) des voyelles orales du Français produites par 10 Tchèques en isolation (marquées par un rond), en contextes labial (triangle), coronal (croix), palato-vélaire (losange) et uvulaire (triangle vers le bas). Moyennes calculées à partir de 4 répétitions* syllabe finale* valeur prise à la moitié de la durée vocalique / 4 répétitions (logatomes) * 3 valeurs : à 1/3, 1/2 et 2/3 de la durée vocalique pour les voyelles isolées. Remarquons une grande proximité acoustique des voyelles moyennes insérées dans le même contexte phonétique

Le chapitre 12 a également montré que la perception et la production des voyelles moyennes est généralement liée, mis à part le cas de la voyelle [ɛ] « hautement similaire » aux [ɛ/ɛ:] tchèques. En effet, cette voyelle est produite authentiquement (avec les formants F1, F2 et F3 à moins d'un écart type de la "référence" et catégorisée comme telle par les natifs) mais elle est identifiée avec difficulté (elle est confondue par les Tchèques avec /e/). Ainsi dans le cas de [ɛ], la performance en production précède la performance en perception et cette situation n'est envisagée ni par le modèle PAM ni par le modèle NLM qui basent les prédictions à propos de l'apprentissage phonétique sur la perception uniquement. A l'inverse, le SLM admet la possibilité que les non-natifs puissent produire certains sons plus ou moins authentiquement sans en posséder la catégorie perceptive.

Pour résumer, la correction phonétique chez les Tchèques du niveau avancé en français devrait viser en priorité les voyelles moyennes dont la distribution n'est pas aléatoire (Wioland, 2005).

De cette manière, les futures enseignantes de FLE doivent être au courant non seulement de la distribution générale mais également des règles particulières notamment dans les phrases où le mauvais choix du timbre pourrait entraîner des erreurs phonologiques, sources de malentendus. En effet comme l'a souligné Madame Vaissière (communication personnelle), il y a une différence entre dire « Pour être de nouveau en forme, je me ferais bien un jeûne (prononcé [ʒøn]) » et « Pour être de nouveau en forme, je me ferais bien un jeune (prononcé [ʒœn]) ». Il reste à vérifier si ce contraste, de même que d'autres contrastes, sont toujours perçus par les jeunes Français non-méridionaux et s'il s'avère donc justifié d'enseigner les règles particulières (contextes oppositifs où les voyelles moyennes ont toujours le statut de phonèmes) aux apprenants du français. Ce point fera l'objet du chapitre 13.

13. Voyelles moyennes du français chez des Français natifs non-méridionaux

Ce que cache mon langage, mon corps le dit. Mon corps est un enfant entêté, mon langage est un adulte très civilisé.

Roland Barthes

Résumé : Par l'analyse des formants F1, F2, (F3), le **chapitre treize** met en évidence le fait que les jeunes Français natifs non-méridionaux et cultivés produisent six qualités vocaliques distinctes correspondant aux voyelles moyennes. Les paires de voyelles e/ɛ et ø/œ se différencient essentiellement par le F1 (et le F2 joue un rôle sur leur authenticité) et les voyelles o/ɔ se distinguent sur le plan des deux formants, F1 et F2.

Le résultat des cinq tests de perception passés par des Français non-méridionaux montre que la distribution des voyelles moyennes obéit avant tout aux règles orthoépiques selon lesquelles /e/ n'apparaît jamais en syllabe finale fermée alors que /ɔ/ et /œ/ n'apparaissent pas en syllabe finale ouverte en français standard. En revanche, les deux couples de voyelles ø/œ et o/ɔ peuvent contraster en syllabe finale ouverte et le couple e/ɛ en syllabe finale fermée. Bien qu'il existe théoriquement de nombreuses paires minimales opposant o/ɔ et e/ɛ, les résultats montrent que les Français non-méridionaux ne sont pas toujours sensibles à l'inversion du timbre dans un grand nombre de contextes oppositifs (l'inversion du timbre ne change pas le sens). Dans le cas où le timbre accepté diffère de celui prescrit par le dictionnaire, il s'avère que la perception des natifs est essentiellement guidée par la loi de position avec une préférence pour les voyelles ouvertes en syllabe fermée et pour les voyelles fermées en syllabe ouverte. La distribution complémentaire s'applique notamment aux voyelles de la paire e/ɛ où [e] est largement préféré au [ɛ] en syllabe finale ouverte. En ce sens, nos résultats appuient notamment le modèle didactique de Wioland (2005) qui recommande aux apprenants de FLE voulant 'faire français' la fermeture du timbre des voyelles moyennes en syllabe finale ouverte.

Le statut phonologique des voyelles moyennes a été confirmé car malgré un faible rendement phonologique des contrastes, il existe des paires minimales où l'inversion du timbre change selon tous les auditeurs le sens de l'unité sémantique. Ces exceptions doivent être présentées à tous ceux qui souhaitent acquérir un niveau phonético-phonologique quasi-natif du français parlé, ce qui devrait être le cas des enseignants de FLE

Le chapitre 12 a pointé la difficulté des Tchèques apprenant le français à identifier et à produire la qualité exacte des six voyelles moyennes que l'on inclut traditionnellement dans le système phonologique français à quatre degrés d'aperture. Il a mis en évidence une confusion des voyelles isolées e/ɛ, o/ɔ et ø/œ.

En théorie, il existe un grand nombre de paires minimales qui portent sur le contraste e/ɛ en syllabe ouverte (Féry, 2011). Il s'agit de quelques morphèmes lexicaux, comme *gré/ grès, foré/ forêt, vallée/ valet*, etc., et de nombreux morphèmes grammaticaux où la qualité vocalique mi-fermée ou mi-ouverte exprime traditionnellement le futur ou le conditionnel, le passé composé ou l'imparfait, le présent indicatif ou le présent subjonctif, comme dans *j'irai/ j'irais, mangé/ mangeais, j'ai/ (que) j'aie* (Léon, 1993). Selon Walter (1976), les oppositions phonémiques se produisent plus systématiquement lorsqu'elles constituent des lexèmes plutôt que des grammèmes mais si l'on en croit Fagyal *et al.* (2002a), elles tendent à disparaître également dans des paires minimales constituées de mots lexicaux. Lefebvre (1988) observe peu de variations dans la réalisation du /e/ des articles possessifs *mes, tes, ses* et des morphèmes lexicaux se terminant par *-é, -ée, -és, -ées, -er* et elle note plus de variations pour les

articles non possessifs *des, les* et pour les terminaisons verbales infinitive *-er* et participiales *-é, -ée, -és, -ées*. De cette façon, le nombre élevé de paires minimales possibles basées sur le contraste *e/ε* ne garantit pas le maintien du contraste en syllabe finale ouverte. Par ailleurs, en syllabe finale fermée, seule la qualité mi-ouverte [ɛ] apparaît en français standard.

Selon Landick (1995), le contraste des voyelles moyennes le plus productif serait celui des voyelles *o/ɔ* qui peuvent s'opposer en syllabe finale fermée. Les contraintes relatives à la distribution des timbres en syllabe CVC en fonction du contexte phonémique sont résumées par Léon (1976) : la voyelle mi-fermée [o] précède toujours la consonne [z], alors que la voyelle mi-ouverte [ɔ] se trouve toujours devant les consonnes dorsales voisées [ɰ, g, ɲ, ŋ] en français standard. En syllabe finale ouverte, seul le timbre mi-fermé [o] apparaît.

Enfin, le contraste assuré par les voyelles *ø/œ* en syllabe finale fermée présente un faible rendement phonologique, car il se produit en théorie dans deux paires minimales uniquement : *jeune/jeûne* et *veulent/veule* (adj). Sa distribution en CVC dépend du contexte phonémique : devant /t, z, ʒ, d/ seule la voyelle /ø/ apparaît alors que devant /R, j, f, v, p, b/, seule la voyelle /œ/ est distribuée en français standard (Léon, 1976). En production, le statut phonémique des /ø, œ/ est instable en français contemporain (Gottfried, 1984) mais pour Detey *et al.* (2010, p. 175), "*ces oppositions semblent ... encore représentatives du standard à "l'oreille" des auditeurs*".

Une tendance vers la distribution complémentaire des voyelles moyennes a été mise en évidence dans les enquêtes phonologiques sur le français dont Martinet est le pionnier (Martinet, 1945). Emprisonné pendant la 2^e guerre mondiale dans un camp d'officiers, il recueillit 409 témoignages sur la prononciation du français. Dans le but de rendre compte du parler spontané et naturel, et de déterminer la « norme » du français, nécessaire pour son enseignement aux étrangers, il posa 45 questions dont douze sur la prononciation des voyelles moyennes. Il classa alors les réponses selon l'âge, en distinguant le groupe des seniors, moyens et juniors, et la région sans confondre les résultats des locuteurs du Midi avec ceux des autres régions de France. La distribution complémentaire ou encore « loi de position » des voyelles moyennes est en effet systématiquement employée dans le sud de la France où les six voyelles moyennes correspondent à trois phonèmes uniquement (Shelly, 1998). Une tendance inverse pour les voyelles mi-ouvertes en syllabe fermée est cependant observée en Lorraine, Alsace et Bretagne où elles tendent à la fermeture. Dans ces régions, le mot <mère> peut se prononcer avec [e] et <peur> avec [ø]. Ainsi, Léon (1993, p. 23) définit les voyelles moyennes comme étant des « *voyelles dont les timbres sont si proches qu'ils sont parfois confondus par des groupes sociaux, géographiques, ou encore dans certains contextes linguistiques.* ».

Du fait de la loi de position qui semble régir de plus en plus la distribution des voyelles moyennes du français, de nombreux enseignants préconisent un modèle pédagogique qui place les voyelles fermées en syllabe ouverte et les voyelles ouvertes en syllabe fermée. En effet, selon Wioland (2005, p. 108), « *Nous sommes [...] devant une évolution du système qui tend à généraliser un seul timbre pour une même structure syllabique.* » et « *Tout l'effort didactique doit porter sur la réalisation fermée des voyelles /o/, /ø/ et /e/ en syllabe CV.* ». De plus, l'attention devrait porter sur le timbre de la syllabe finale du mot prosodique où la qualité est pleinement réalisée. En syllabe non-finale (et donc inaccentuée), le timbre de la voyelle serait plutôt légèrement ouvert mais d'un point de vue didactique il conviendrait de le transcrire par l'archiphonème /E/, /œ/ et /O/ pour ne pas attirer inutilement l'attention des apprenants. A ce propos, Asla Ostby (2005) précise que les voyelles moyennes en position faible peuvent subir le processus de la dilation qui est selon Fagyal *et al.* (2002b) le plus productif dans les formes dérivées. Ainsi, la qualité vocalique de la syllabe inaccentuée est influencée par la voyelle de la syllabe accentuée comme dans les exemples suivants : [ɛ] dans le mot *aimable* (Asla Ostby, 2005), [ɔ] dans le mot *dossard* (Morin, 1983) et [œ] dans le mot *aveuglant* (Fagyal *et al.*, 2002b).

Une autre position vis-à-vis l'enseignement des voyelles du français est celle de Shelly (1998) qui considère que le contraste devrait être présenté et enseigné aux apprenants avancés, susceptibles de

le rencontrer dans de nombreuses variétés du français. Ainsi en fonction du public visé, les enseignants peuvent adopter une approche *minimaliste* ou *maximaliste* (Galazzi-Matasci and Pedoya, 1983).

L'approche dans la présente thèse est naturellement *maximaliste* car nous visons les futurs enseignants de FLE, censés posséder une performance quasi-native de la langue enseignée (Leather, 1983). Cependant ne voulant en aucun cas imposer des règles obsolètes, nous nous intéressons dans ce chapitre à la sensibilité des Français natifs non-méridionaux aux contrastes entre les voyelles moyennes.

13.1 Tests de perception

Dans le but d'examiner quelles sont les règles de distribution des voyelles moyennes du français chez les jeunes Français non-méridionaux et leur statut phonémique, nous avons effectué une série de cinq tests perceptifs qui sont les suivants :

1. Tâche de la singularité²⁴ (test de discrimination)
2. Identification de mots prononcés comme en français standard
3. Identification de mots dans des paires minimales
4. Identification de phrases avec sens inattendu
5. Identification de sons vocaliques générés avec la synthèse de Klatt

Les quatre premières études sur la perception des voyelles moyennes du français ont été effectuées par cinq Français natifs (un homme et quatre femmes) non-méridionaux ayant longtemps habité en région parisienne et grandi à Gien (auditeur 1), Lyon (auditrice 2), Paris (auditrice 3), Enghien les Bains (auditrice 4) et St. Briec (auditrice 5). Ils ont tous été étudiants ou enseignants à Paris 3 au moment des tests. Les informations relatives aux cinq Français natifs se trouvent dans le Tableau 116.

Français natif	Lieu où le natif a grandi	Habitant la région parisienne depuis (nombre d'années)	Age (J- junior < 20 ans, M- moyen < 50 ans, S- senior 50 ans et plus)
1	Gien (24)	4 ans	M
2	Lyon (69)	8 ans	M
3	Paris (75)	Toujours	M
4	Enghien les Bains (95)	Toujours	M
5	St. Briec (22)	5 ans	M

Tableau 116 : Informations sur les 5 natifs du français ayant répondu au questionnaire phonologique et effectué les tests de perception sur les voyelles moyennes du français

La cinquième étude a été effectuée par un auditeur français natif, âgé de 37 ans, ayant grandi dans la région des Pays de la Loire et habité 15 ans à Paris.

²⁴ En anglais « oddity task »

13.1.1 Tâche de la singularité

La construction et les conditions de passation du premier test perceptif portant sur les voyelles moyennes est indiqué dans le Tableau 117. Il s'agit d'une tâche de la singularité où les cinq auditeurs français natifs ont écouté 48 suites de quatre monosyllabes de type CV pour le contraste e/ɛ et de type CVC pour les contrastes o/ɔ et ø/œ, où C correspondait à l'une des consonnes [p, t, k, ʁ]. Les quatre monosyllabes (qui sont des non-mots) de chaque suite avaient été enregistrés par des locutrices françaises natives non-méridionales différentes et les auditeurs indiquaient si les voyelles contenues dans les quatre monosyllabes de chaque suite étaient identiques. Le cas échéant, ils indiquaient si la voyelle qui différait des trois autres se trouvait dans la 1^e, 2^e, 3^e ou 4^e syllabe. Le test commence par dix suites d'entraînement et contient trois suites de contrôle qui sont composées des voyelles identiques.

Type de test	De discrimination
Stimuli	3 paires de voyelles moyennes dans des syllabes CV(C) où C = [p, t, k, ʁ], 16 suites pour chaque paire examinée, exemple d'une suite : ke pɛ te Re
Consigne	Les quatre syllabes de chaque suite contiennent-elles la même voyelle ? Cochez oui ou non. S'il y a une voyelle différente des autres, dites dans quel mot elle se trouve : 1 ^e , 2 ^e , 3 ^e ou 4 ^e mot de la suite ?
Auditeurs	5 Français natifs non méridionaux
Locuteurs	1 Français natif (Pays de la Loire)
Nb total de stimuli	61
Nb de stimuli d'entraînement	10
Nb de stimuli distracteurs	3

Tableau 117 : Information sur la construction et les conditions de passation du test de discrimination des voyelles moyennes du français par 5 Français natifs non méridionaux

Les suites de monosyllabes proposées aux auditeurs se trouvent dans le Tableau 118. Les suites grisées correspondent aux stimuli d'entraînement ou de contrôle dont les réponses n'ont pas été incluses dans les résultats.

1.	pøp töt kæk RøR	21.	töt RøR kæk pøp	41.	køk pøp töt RøR
2.	töt køk pøp RøR	22.	Rε pe ke te	42.	Rε te pe ke
3.	RæR töt pøp køk	23.	ke pe te Re	43.	tøt køk pøp RæR
4.	pop tot køk RøR	24.	pøp töt køk RæR	44.	töt køk RøR pop
5.	pøp køk RøR töt	25.	töt køk pøp RøR	45.	pe te ke Rε
6.	tot pop køk RøR	26.	pe te ke Rε	46.	kæk RæR pøp töt
7.	te pe ke Rε	27.	pøp töt køk RøR	47.	køk RøR pop töt
8.	te ke pe Re	28.	töt køk RøR pop	48.	pε te ke Rε
9.	pε te ke Re	29.	ke pε te Re	49.	pop tot køk RøR
10.	pop tot køk RøR	30.	tøt kæk RøR pøp	50.	køk pøp töt RæR
11.	pøp kæk töt RøR	31.	pøp køk RøR töt	51.	pop tot køk RøR
12.	pop tot køk RøR	32.	pε ke Re te	52.	pε Re te ke
13.	te ke Re pe	33.	RæR køk pøp töt	53.	töt pøp køk RøR
14.	pøp køk RøR töt	34.	pøp RøR töt køk	54.	pøp tot køk RøR
15.	töt pøp RøR køk	35.	te Rε ke pε	55.	Re ke te pe
16.	ke te Re pe	36.	pøp töt køk RøR	56.	RæR kæk pøp töt
17.	kæk RæR pøp töt	37.	töt RøR kæk pøp	57.	töt køk RøR pøp
18.	töt RøR pøp køk	38.	töt køk pop RøR	58.	pε ke te Rε
19.	ke te re pε	39.	te Re pε ke	59.	RæR töt kæk pøp
20.	RøR kæk pøp töt	40.	pøp køk töt RøR	60.	køk tot pøp RøR
61.	Re ke pε te				

Tableau 118 : Suites de mots proposées aux 5 auditeurs Français natifs non méridionaux dans la tâche de la singularité. Les suites grisées sont des stimuli d'entraînement ou de contrôle

Le résultat du test qui indique le nombre de fois où les voyelles ont été “correctement” discriminées par chaque auditeur séparément se trouve dans le Tableau 119. Alors que les auditeurs 1, 2 et 5 ont “correctement” discriminé toutes les voyelles présentées (100 %) de réussite), l’auditeur 3 n’a pas perçu le contraste e/ε dans une suite, le contraste o/ɔ dans deux suites et le contraste ø/œ dans 4 suites. Enfin, l’auditeur 4 n’a pas perçu le contraste ø/œ dans une suite.

	e/ɛ sur 16	o/ɔ sur 16	ø/œ sur 16
1 : Gien	16	16	16
2 : Lyon	16	16	16
3 : Paris	15	14	12
4 : Enghien les Bains	16	16	15
5 : St. Briec	16	16	16
TOTAL	79/80 (99%)	78/80 (98%)	75/80 (94%)

Tableau 119 : Résultat de la tâche de la singularité sur la discrimination des voyelles moyennes dans des suites de 4 mots, effectuée par 5 auditeurs français natifs non méridionaux (leur région d'origine est indiquée en colonne 1)

Ainsi, dans la tâche de la singularité, la productivité des contrastes baisse de e/ɛ (confusion globale de 1 % : le contraste n'est pas perçu dans une suite sur 80) à o/ɔ (confusion globale de 2 % : le contraste n'est pas perçu dans deux suites sur 80) à ø/œ (confusion globale de 6 % : le contraste n'est pas perçu dans cinq suites sur 80).

Globalement, la tâche de la singularité indique une bonne discrimination des voyelles moyennes par des auditeurs français non-méridionaux natifs.

13.1.2 Identification du français standard

Dans le deuxième test perceptif, les cinq auditeurs français natifs non méridionaux ont écouté 24 phrases et indiquaient si le dernier mot de chaque phrase était prononcé comme en français standard (sans accent régional). L'accent régional perçu pouvait provenir de l'inversion du timbre au sein des trois paires des voyelles moyennes, insérées dans des syllabes accentuées en fin de groupes de sens.

Type de test	(D')identification
Stimuli	Phrases avec et sans inversion de timbre des voyelles moyennes en syllabe finale
Consigne	Le dernier mot est-il prononcé comme en français standard?
Auditeurs	5 Français natifs non méridionaux
Locuteurs	1 Français natif (Pays de la Loire)
Nb total de stimuli	24
Nb de stimuli d'entraînement	0
Nb de stimuli distracteurs	12

Tableau 120 : Information sur la construction et les conditions de passation du test d'identification des voyelles moyennes du français par 5 Français natifs non méridionaux

Ainsi, huit phrases (deux fois quatre) avaient été prononcées avec [ɔ] et [œ] en syllabe finale CV et quatre phrases avec [e] en syllabe finale CVC ce qui enfreint les règles orthoépiques du français. En revanche, douze phrases avaient été prononcées selon l'indication du dictionnaire²⁵. La construction du test ainsi que les conditions de passation sont décrites dans le Tableau 127.

Les 24 phrases que nous avons fait écouter aux auditeurs français se trouvent dans le Tableau 121. Les mots contenant les voyelles moyennes qui peuvent engendrer la perception d'un accent régional sont marqués en gras.

1.	Prends-les.
2.	Tu aimes ce jeu ? (prononcé avec [œ])
3.	Cet été, on va à la mer.
4.	Achète du sel . (prononcé avec [e])
5.	Sur la chaussure, on fait un nœud.
6.	Il est là- haut . (prononcé avec [ɔ])
7.	Il a un bouclier et une épée.
8.	Tu peux le faire ? Oui, je peux . (prononcé avec [œ])
9.	Le trou est foré.
10.	Attention, le café est chaud . (prononcé avec [ɔ])
11.	Qu'est-ce que ça vaut ?
12.	La femme est belle . (prononcé avec [e])
13.	Il fait beau . (prononcé avec [ɔ])
14.	Il y est allé de son gré.
15.	Elle a une belle peau.
16.	Je te présente mon père . (prononcé avec [e])
17.	Ce garçon est beau.
18.	<u>J'ai des bonbons, tu en veux ?</u> (prononcé avec [œ])
19.	Il fait très chaud.
20.	Je voudrais une bouteille d'eau . (prononcé avec [ɔ])
21.	Il faut faire la queue.
22.	Tu as des enfants ? Oui, j'en ai deux . (prononcé avec [œ])
23.	Qu'est-ce qu'elle est belle !
24.	Au printemps, je prends ma fille à la mer . (prononcé avec [e])

Tableau 121 : 24 phrases proposées aux 5 auditeurs Français natifs non méridionaux qui devaient identifier les phrases prononcées avec un accent régional pouvant être induit par l'inversion du timbre au sein des paires de voyelles moyennes (mots en gras). La phrase où l'inversion n'a pas été identifiée par l'auditeur 2 est soulignée

Le Tableau 121 indique le pourcentage d'identification de l'inversion du timbre au sein des paires de voyelles moyennes par les auditeurs un à cinq et les résultats détaillés se trouvent en annexe du chapitre 13, à la page 89 des annexes.

Le résultat du Tableau 122 indique que l'inversion du timbre au profit de la qualité mi-fermée [e] en syllabe finale CVC et de la qualité mi-ouverte [ɔ] en syllabe finale CV a été identifiée dans 20 réponses sur 20 (100 %) comme du français non-standard. Puis, l'inversion des voyelles ø/œ au profit de la qualité mi-ouverte en syllabe finale CV a été identifiée comme du français non-standard dans 19 réponses sur 20, soit la quasi totalité. Ce résultat indique que la perception des auditeurs Français natifs non-méridionaux est régie par les règles orthoépiques du français selon lesquelles /e/ n'apparaît jamais en syllabe finale fermée et /ɔ, œ/ jamais en syllabe finale ouverte quoique une auditrice ait jugé la prononciation [vœ] du mot *veux* comme standard.

²⁵ www.lerobert.com

Identification de l'inversion	e/ɛ sur 4	o/ɔ sur 4	ø/œ sur 4
1 : Gien	4	4	4
2 : Lyon	4	4	3
3 : Paris	4	4	4
4 : Enghien les Bains	4	4	4
5 : St. Briec	4	4	4
TOTAL	20/ 20	20/ 20	19/ 20

Tableau 122 : Résultat de la tâche d'identification par 5 Français non-méridionaux de 4 phrases prononcées avec [e] en syllabe CVC, 4 avec [ɔ] et 4 avec [œ] en syllabe CV

Ainsi, les règles orthoépiques sur la distribution semi-complémentaire des voyelles moyennes en fonction de la structure syllabique sont dans notre test activement employées par des Français natifs non-méridionaux et déterminent la prononciation jugée comme français standard.

13.1.3 Identification de mots dans des paires minimales

Dans la troisième tâche perceptive, les cinq auditeurs français natifs des tests précédents ont identifié 61 mots prononcés par un Français natif non méridional en entourant l'un ou l'autre mot constituant une paire minimale. Le contraste reposait de nouveau sur les voyelles moyennes e/ɛ, o/ɔ, ø/œ. La construction du test ainsi que les conditions de passation sont décrites dans le Tableau 123

Type de test	(D')identification
Stimuli	Mots des paires minimales basées sur le contraste entre les voyelles moyennes
Consigne	Entourez le(s) mot(s) que vous entendez
Auditeurs	5 Français natifs non méridionaux
Locuteurs	1 Français natif (Pays de la Loire)
Nb total de stimuli	61
Nb de stimuli d'entraînement	0
Nb de stimuli distracteurs	0

Tableau 123 : Information sur la construction et les conditions de passation du test d'identification par 5 Français natifs non méridionaux des mots dont le contraste est basé sur la qualité des voyelles moyennes du français

Les paires minimales proposées se trouvent dans le Tableau 124. Trente-et-une paires minimales y sont basées sur le contraste o/ɔ, 26 sur le contraste e/ɛ et quatre sur ø/œ. La qualité effectivement prononcée est en gras.

13. Voyelles moyennes du français chez des Français natifs non-méridionaux

1.	Paume - pomme	21.	Veule (adj.) - veulent	41.	Beauce - bosse
2.	Blet - blé	22.	Taupe - top	42.	Mais - mes
3.	Les - lait	23.	Saule - sol	43.	Rauque - rock
4.	Taie - thé	24.	Vautre - votre	44.	Valet - vallée
5.	J'aurai – j'aurais	25.	Veulent – veule (adj.)	45.	Rauque - rock
6.	Craie - créé	26.	Mais - mes	46.	Dé - dès
7.	Blé - blet	27.	Saule - sol	47.	Baie - B
8.	Homme - heaume	28.	Valet - vallée	48.	Etait - été
9.	Les - lait	29.	Sonne – Saône	49.	Beauce - bosse
10.	Manger - mangeais	30.	Vautre - votre	50.	Cote – côte
11.	Col – khôl	31.	Aude - ode	51.	Tournaient - tournée
12.	Thé - taie	32.	Col – khôl	52.	Paul - Paule
13.	Craie - crée	33.	Os - hausse	53.	Molle – môle
14.	Sonne – Saône	34.	Taupe - top	54.	Dé - dès
15.	J'aurais – j'aurai	35.	Aude - ode	55.	Maur - mort
16.	Mangeais - manger	36.	Paul – Paule – pôle	56.	Molle – môle
17.	Fée - fait	37.	Jeûne - jeune	57.	Pôle – Paul - Paule
18.	Paume - pomme	38.	Hausse - os	58.	Baie - B
19.	Homme - heaume	39.	Fée - fait	59.	Cote – côte
20.	Jeune – jeûne	40.	Pôle – Paul	60.	Etait - été
61.	Tournaient - tournée				

Tableau 124 : Paires minimales proposées dans le test d'identification des mots contenant une voyelle moyenne. La qualité effectivement prononcée est en gras

Le résultat du dernier test de perception (voir Tableau 125 page 284) montre que selon 144 réponses sur 155 (soit 93 %), les auditeurs français natifs non méridionaux sont perceptivement sensibles au contraste o/ɔ, selon 117 réponses sur 130 (soit 90 %) au contraste e/ɛ de même que selon 18 réponses sur 20 (soit 90 %) au contraste ø/œ.

Identification	e/ɛ sur 26	o/ɔ sur 31	ø/œ sur 4
1 : Gien	23	29	4
2 : Lyon	25	31	4
3 : Paris	23	28	3
4 : Enghien les Bains	22	27	3
5 : St. Briec	24	29	4
TOTAL	117/ 130 (90%)	144/ 155 (93%)	18/ 20 (90%)

Tableau 125 : Résultat du test d'identification par 5 Français natifs non méridionaux de 26 mots basés sur le contraste e/ɛ, 31 mots sur o/ɔ et 4 mots sur ø/œ

En analysant plus en détail ces réponses (qui se trouvent intégralement en annexe du chapitre 13, à la page 89 des annexes), nous remarquons que les paires minimales au sein desquelles le contraste est perçu par les auditeurs à l'unanimité sont les suivantes :

Contraste	Paires minimales
e/ɛ	Blet – blé Taie – thé Craie – crée Mangeais - manger Fait - fée Dès - dé Baie – B
o/ɔ	Paume – pomme Heaume - homme Taupe – top Saulé – sol Saône - sonne Vautre – votre Beauce – bosse Rauque – rock Môle - molle
ø/œ	Veule (adj.) - veulent

Tableau 126 : Liste des paires minimales dans lesquelles les contrastes e/ɛ, o/ɔ et ø/œ sont systématiquement perçus par 5 auditeurs français natifs non-méridionaux

Selon les résultats, l'opposition o/ɔ est la plus faible dans la paire *Aude* – *ode* (mots confondus par 3 auditeurs), puis dans *Paul* – *pôle* – *Paule*, *cote* – *côte* (confondus par 2 locuteurs) et enfin dans *col* – *khôl* et *os* – *hausse* (confondus par un locuteur). La productivité du contraste ø/œ, testée à travers la perception de mots apparaissant dans seulement deux paires minimales n'est pas systématique pour *jeune* – *jeûne* car les mots sont confondus par deux auditeurs. Enfin, la confusion des timbres e/ɛ est

maximale quand le contraste est contenu dans des grammèmes, comme c'est le cas des paires *j'aurai* – *j'aurais* (confondus par 4 auditeurs), *mes* – *mais* (confondus par 3 auditeurs), *les* – *lait* (confondus par 2 auditeurs) et *était* – *été* de même que *tournaient* – *ournée* (confondus par un auditeur). Les seuls lexèmes confondus étaient ceux de la paire *valet* – *vallée* (confondus par un auditeur).

Dans l'analyse des réponses, il est important de préciser que **sur les treize confusions de qualités e/ɛ, dix confusions étaient de type ɛ -> e et seulement trois de type e -> ɛ. La préférence du timbre fermé en syllabe finale ouverte** montre que lorsque le contraste entre les voyelles e/ɛ n'est pas perçu, ce qui arrive notamment quand ces dernières se trouvent dans des grammèmes, la **perception des auditeurs est largement guidée par la loi de position**.

Or ce n'est plus le cas dans le cas du contraste ø/œ où deux auditeurs relient le timbre [ø] au mot *jeune*, qui devrait être produit selon le dictionnaire²⁶ et en même temps selon la loi de position avec [œ]. Enfin, la perception des cinq auditeurs dans ce test n'est pas non plus guidée par la distribution complémentaire des timbres o/ɔ car sur les onze confusions, cinq confusions sont de type ɔ -> o visant donc la fermeture en syllabe finale fermée.

Le résultat de ce test perceptif a confirmé le statut phonémique des voyelles moyennes car dans certaines paires minimales, l'un ou l'autre qualité renvoie systématiquement à l'un ou l'autre mot de la paire. Nous avons également montré que dans les contextes oppositifs où le contraste e/ɛ n'est plus actif, la perception des auditeurs est guidée par la loi de position.

13.1.4 Identification de sens « inattendu »

Dans la quatrième tâche perceptive, les mêmes cinq auditeurs français natifs non méridionaux ont écouté 58 phrases et indiquaient si les phrases avaient un sens inattendu (bizarre, drôle). Le sens inattendu provenait éventuellement de l'inversion du timbre au sein des trois paires de voyelles moyennes, insérées dans des syllabes accentuées en fin de groupes de sens.

Ainsi, seize phrases avaient été prononcées avec l'inversion des voyelles o/ɔ (dont sept au profit de [o] et neuf du [ɔ]), onze avec l'inversion des voyelles e/ɛ (dont cinq au profit du [e] et six du [ɛ]) et quatre avec l'inversion de ø/œ (dont deux au profit du [ø] et deux du [œ]). En revanche, 27 phrases avaient été prononcées selon les règles particulières de prononciation indiquées dans le dictionnaire²⁷. Les détails concernant ce test sont décrits dans le Tableau 127.

Type de test	(D')identification
Stimuli	Phrases sans et avec inversion du timbre des voyelles moyennes en syllabe finale
Consigne	Le sens de la phrase est-il inattendu ?
Auditeurs	5 Français natifs non méridionaux
Locuteurs	1 Français natif (Pays de la Loire)
Nb total de stimuli	58
Nb de stimuli d'entraînement	0
Nb de stimuli distracteurs	27

Tableau 127 : Information sur la construction et les conditions de passation du test d'identification des voyelles moyennes du français par 5 Français natifs non méridionaux

²⁶ www.lerobert.com

²⁷ www.lerobert.com

Les 58 phrases que nous avons fait écouter aux auditeurs français se trouvent dans le Tableau 128. Les mots contenant les voyelles moyennes qui peuvent engendrer un sens inattendu du fait de l'inversion du timbre y sont marqués en gras et les phrases ayant été effectivement identifiées comme inattendues à l'unanimité par tous les cinq auditeurs sont soulignées. Notons que nous avons testé la perception de l'inversion du timbre uniquement dans des mots lexicaux.

1. Les impôts sont en hausse.
2. Pour être de nouveau en forme, je me ferais bien un **jeune**.
3. Au Moyen-âge, les chevaliers portaient une épée, un bouclier et un **homme**.
4. Tu pourras venir ? Oui, je pourrai.
5. Il fait froid. Ferme ton **khôl**.
6. Pierre aime Aude.
7. Le chômage est en **os**.
8. Tu veux du café ou du **taie** ?
9. Il faut lancer les dés.
10. Il tient une coccinelle dans sa **pomme**.
11. Mon bébé boit encore du lait.
12. Il s'appelle **Paule**.
13. Il fait beau.
14. J'aime le jazz et le rock.
15. En France, j'aime en particulier **la bosse**.
16. On a acheté un nouveau lit : il faut encore des draps et des **thés**.
17. Tu as déjà mangé ?
18. Dans le magasin, ils vendent des jupes et des **taupes**.
19. Le roi avait des serviteurs et des **vallées**.
20. Il est gentil : j'aime cet **heume**.
21. Alain Souchon est en tournée.
22. Je dois acheter deux poires et une **paume**.
23. Le long de l'océan Atlantique, il y a plusieurs **B**.
24. Sur la chaussure, on fait un nœud.
25. Pour trouver le livre dans le rayon, il faut connaître sa **côte**.
26. Ma mère est encore jeune.
27. Les enfants à l'école savent déjà dessiner les A, maintenant ils apprennent à dessiner les **baies**.
28. La poire est trop molle.
29. Ma sœur s'appelle **Paul**.
30. Mes initiales sont B. P.
31. Pour se maquiller, on utilise le **col**.
32. Ca sonne, va ouvrir vite !
33. Il l'a fait contre son **grès**.
34. Dans la vallée il fait plus chaud qu'en haut de la montagne, c'est sûr.
35. La France compte 65 millions d'habitants et 27 régions, ce sont des **fées**.
36. J'aime votre fille.
37. Tu dois leur acheter ces livres : elle est **veule**.
38. Je voudrais du sucre dans mon thé.
39. Autrefois, l'Afrique du nord était habitée par des **morts**.
40. Le jardin est dévasté par des **top**.
41. Les hommes sont partis.
42. En l'honneur des dieux, on chantait des **Audes**.
43. Paul et Marie s'aiment.
44. La bosse est grande : je me suis vraiment bien tapé.
45. Mes chats sont dehors.

46.	Sa voix est enrouée et rock .
47.	Au pied de la montagne, il y a plusieurs valets .
48.	Il fait froid dehors, ferme ton col !
49.	<u>Il faut avoir un six dans un jeu de dès.</u>
50.	Mon chien est mort.
51.	<u>Dans les comptes pour enfants, il y a des princesses et des faits.</u>
52.	Il y a une taupe dans le jardin.
53.	J'aime les fraises et les pommes.
54.	<u>Je le trouve déprimé : même le psychologue dit qu'il (l)es veulent.</u>
55.	Il était sur place.
56.	Au-dessus du lit de la reine, il y a un dé .
57.	La côte vendéenne est belle.
58.	<u>Quand on aime les jeûnes, on va à la discothèque</u>

Tableau 128 : 58 phrases proposées aux 5 auditeurs Français natifs non méridionaux qui ont identifié les phrases ayant un sens inattendu (bizarre, drôle) induit par l'inversion du timbre au sein des paires des voyelles moyennes (mots en gras). Les phrases dont le sens est inattendu selon tous les 5 auditeurs sont soulignées

Le Tableau 129 indique le pourcentage d'identification de l'inversion du timbre au sein des paires de voyelles moyennes (et donc de l'identification du sens inattendu) par les auditeurs, de un à cinq. Notons que les phrases prononcées selon les règles particulières du dictionnaire (www.lerobert.com) ont été systématiquement identifiées comme non porteuses de sens inattendu.

Identification de l'inversion	e/ɛ sur 11	o/ɔ sur 16	ø/œ sur 4
1 : Gien	7	15	3
2 : Lyon	7	13	3
3 : Paris	7	13	4
4 : Enghien les Bains	7	14	4
5 : St. Brieuc	11	14	4
TOTAL	39/ 55 (71 %)	69/ 80 (86 %)	18/ 20 (90 %)

Tableau 129 : Résultat de la tâche d'identification de 31 phrases bizarres effectuée par 5 auditeurs français natifs non méridionaux

L'analyse du Tableau 129 nous permet les constats suivants :

- le contraste ø/œ est le plus productif chez les cinq auditeurs français non-méridionaux car l'inversion du timbre induisant un sens inattendu a été identifiée dans 18 phrases sur les 20 proposées. Cette « performance » peut s'expliquer par le fait qu'en français moderne l'opposition ø/œ est réalisable que dans deux paires minimales,
- l'inversion du timbre o/ɔ a été identifiée dans 69 réponses sur 80 (soit 86 %). Le nombre relativement plus important de phrases proposées s'explique par l'existence de nombreux contextes véritablement oppositifs,
- le contraste le plus « fragile » est celui des e/ɛ car l'inversion du timbre au sein de cette paire a été identifiée dans 39 réponses sur 55 (soit 71 %). Ce résultat confirme que le manque de sensibilité à l'inversion du timbre au sein de cette paire ne concerne pas seulement les mots grammaticaux, comme le met en évidence Walter (1976) mais également les mots lexicaux comme l'ont démontré Fagyal *et al.* (2002a).

Une analyse approfondie des réponses obtenues (dont l'intégralité est donnée en annexe du chapitre 13, page 90 des annexes) et compte tenu de la prescription du dictionnaire (www.lerobert.com), nous permet d'avancer les remarques suivantes quant à l'inversion des timbres vocaliques :

- ø/œ : la fermeture du timbre en syllabe finale CVC (œ->ø) a été toujours identifiée (induisant un sens bizarre) alors que son ouverture n'a pas été remarquée dans deux cas sur quatre ;
- o/ɔ : la fermeture du timbre en syllabe finale CVC (ɔ->o) a été identifiée dans 32 réponses sur 35 (soit 91 %) alors que son ouverture a été identifiée dans 37 réponses sur 45 (soit 82 %), donc moins souvent ;
- e/ɛ : l'ouverture du timbre en syllabe finale CV (e->ɛ) a été identifiée dans 29 réponses sur 30 (soit 97 %), donc un pourcentage élevé, alors que sa fermeture a été souvent inaperçue car identifiée dans seulement 9 réponses sur 25 (soit 36 %)

Ainsi, pour les timbres dont l'inversion n'a pas été perçue, les réponses dans ce test indiquent que la perception des auditeurs français natifs est généralement régie par la loi de position qui tend vers l'ouverture de la voyelle en syllabe finale fermée (et donc la préférence des timbres [œ, ɔ]) et sa fermeture en syllabe finale ouverte (et donc une préférence pour le timbre [e]).

L'analyse des résultats obtenus montre que l'identification de l'inversion du timbre varie selon les auditeurs et les mots proposés. Il existe néanmoins des mots où la réalisation de la qualité prescrite par le dictionnaire²⁸ (et qui ne suit pas la distribution complémentaire) est attendue à l'unanimité par tous les auditeurs. Ainsi, les mots *heaume*, *hausse*, *paume*, *Beauce* et *taupe* doivent être produits avec la qualité mi-fermée [o] et le mot *veule* (adj.) avec la qualité mi-fermée [ø]. En revanche, nous n'avons relevé aucun mot lexical où la substitution de la qualité mi-ouverte [ɛ] en syllabe finale ouverte (qu'impose le dictionnaire) par la qualité mi-fermée [e] aurait pu surprendre tous nos locuteurs français.

Les futures enseignantes du français, supposées posséder une performance quasi-native en français, devraient connaître ces cas particuliers de prononciation.

13.1.5 Identification de sons vocaliques synthétisés (synthèse à formants de Klatt)

Dans le dernier test perceptif, nous nous sommes intéressée à l'identification de sons vocaliques construits dans Praat avec le synthétiseur à formants de (Klatt, 1980). Les sons générés avaient les caractéristiques acoustiques suivantes :

- Avec une fréquence fondamentale constante de 125 Hz, une durée de 300 ms et les valeurs de F3, F4, F5 et F6 fixés à respectivement 3500, 4500, 5500 et 6500 Hz, 120 sons vocaliques variaient par la valeur du F1 et du F2.
- Le premier formant variait de 340 Hz à 700 Hz par pas de 40 Hz (au total 10 conditions pour le F1).
- Le deuxième formant s'étendait de 700 Hz à 2800 Hz. Les valeurs de F2 ont été espacées de la façon suivante : entre 700 et 1200 Hz les sons ont été générés par pas de 100 Hz, puis par pas de 200 Hz jusqu'à 2000 Hz, puis par pas de 400 Hz jusqu'à 1800 avec un total de 12 conditions pour le F2.

Les 120 sons vocaliques synthétisés sont représentés dans l'espace acoustique F1/F2 (en Hz) de la Figure 113 par des carrés distincts. Les voyelles [i, a, u] représentées servent de point de repère

²⁸ www.lerobert.com

13. Voyelles moyennes du français chez des Français natifs non-méridionaux

(leurs valeurs formantiques F1 et F2 moyennes sont celles calculées en 8.1 à partir de nos données sur les voyelles en isolation produites par 10 Françaises natives* 4 répétitions).

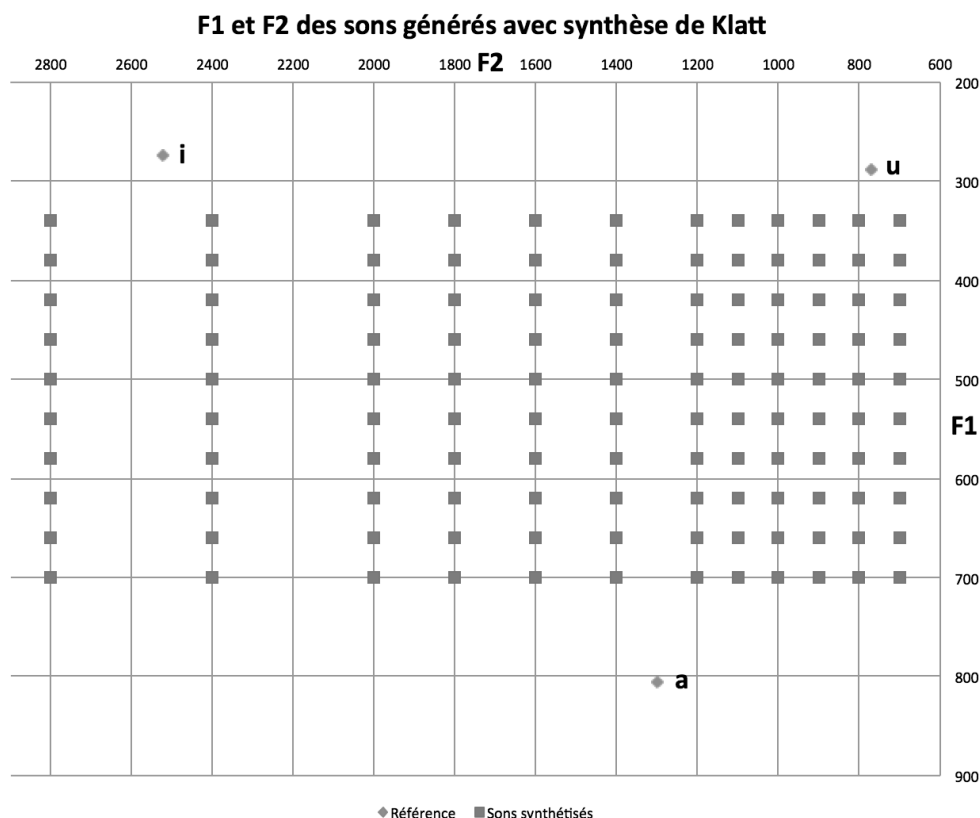


Figure 113 : F1 et F2 des 120 sons vocaliques synthétisés avec Klatt (carrés rouge). Les voyelles [i, a, u] y servent de point de repère

La construction du test de perception est ensuite récapitulée dans le Tableau 130.

Type de test	(D')Identification
Stimuli	120 sons vocaliques synthétisés avec Klatt : F1 varie de 340-700 Hz, F2 de 700-2800 Hz. F0 à 125 Hz, durée 300 ms. F3, F4, F5 et F6 constants à 3500, 4500, 5500 et 6500 Hz
Auditeurs	1 Français natifs, 11 répétitions
Nb total de stimuli	129
Nb de stimuli d'entraînement	9
Nb de stimuli distracteurs	0
Echelle de justesse	1 – 5 1 pas sûr(e) du tout 3 moyennement sûr(e) 5 sûr(e)

Tableau 130 : Construction et passation du test d'identification des sons vocaliques synthétisés avec Klatt

Ce test, contenant au total 129 stimuli (9 stimuli de la phase d'entraînement et 120 stimuli du test) a été passé onze fois (à un jour d'intervalle au minimum) par un auditeur français natif de 37 ans, éduqué, ayant grandi dans la région des Pays de la Loire et ayant vécu 15 ans à Paris. Cet auditeur a déjà passé les tests précédents portant sur l'identification des voyelles du français où il s'est avéré qu'il suit la tendance générale (il est donc représentatif des auditeurs français actuels) et qu'il exploite au maximum l'échelle exprimant le niveau de confiance.

L'interface graphique du test indiquant la consigne et les dix boutons de réponse est illustrée à la Figure 114. Les mots proposés sont « ville » pour [i], « sûr » pour [y], « bébé » pour [e], « mer » pour [ɛ], « peu » pour [ø], « fleur » pour [œ], « papa » pour [a], « où » pour [u], « photo » pour [o] et « robe » pour [ɔ]. Le niveau de confiance de l'auditeur était quantifié sur une échelle de 1 à 5 où 1 correspondait à « Pas sûr(e) du tout », 5 à « Sûr(e) » et 3 à « Moyennement sûr(e) ».

Figure 114 : Interface graphique du test d'identification des sons vocaliques synthétisés avec Klatt

Les stimuli ayant été identifiés comme /e/ ou /ɛ/ se trouvent dans le Tableau 131 qui indique à quelles fréquences de F1 et F2 ces derniers correspondent et donc sur quel indice acoustique le contraste e/ɛ repose. À côté de chaque catégorie se trouve le *fit index* de l'identification, calculé sur la base de la proportion de l'identification (avec au maximum 11 réponses) et de la note moyenne.

Les résultats du Tableau 131 montrent que le contraste e/ɛ repose essentiellement sur la valeur de F1 : avec F2 entre 2000 Hz et 2800 Hz, les stimuli ayant F1 entre 340 Hz et 500 Hz sont interprétés en termes de /e/ et ceux ayant F1 entre 580 Hz et 700 Hz sont interprétés en termes de /ɛ/. En revanche, lorsque F1 est de 540 Hz (et le F2 entre 2400 Hz et 2800 Hz), les sons renvoient aux deux catégories /e/ et /ɛ/. La valeur de F2 joue ensuite sur la qualité du son perçu : lorsque F2 est élevé (de 2800 Hz), les stimuli catégorisés comme /e/ obtiennent la plus grande note de qualité. En revanche F2 plus bas (à 2400 Hz) favorise la note de qualité maximale pour les sons catégorisés comme /ɛ/. Ensuite lorsque F2 est à 2000 Hz, l'identification en termes de /e/ (avec F1 inférieur à 500 Hz) est très rare alors que celle en termes de /ɛ/ (avec F1 supérieur à 540 Hz) se produit plus souvent quoique avec une note de qualité moyenne plus basse que lorsque F2 se trouve entre 2400 Hz et 2800 Hz. Dans ce sens, nous rejoignons Léon and Tennant (1990) qui définissent la voyelle [e] comme acoustiquement diffuse et la voyelle [ɛ] comme compacte.

13. Voyelles moyennes du français chez des Français natifs non-méridionaux

F2 \ F1	2800	2400	2000	1800
340	e0,9	e2,4		
380	e5	e4,5		
420	e5	e4,4	e0,4	
460	e3,18	e4,3	e0,1	
500	e3,5	e3,9		e0,1
540	e1,9	e1,4		
	<i>ε0,5</i>	<i>ε1,8</i>		
580	ε3,4	ε3,5	ε3,7	
620	ε4,5	ε4,7	ε2,8	ε0,3
660	ε4,5	ε4,8	ε2	
700	ε4,6	ε4,9	ε3,6	

Tableau 131 : Identification des voyelles synthétisées en termes de /e/ ou /ε/ (en italique) avec la valeur du *fit index* de l'identification, calculé sur la base de la proportion de l'identification et de la note moyenne obtenue

Les stimuli répertoriés comme /o/ ou /ɔ/ se trouvent dans le Tableau 132 qui précise leurs fréquences de F1 et F2. La valeur du *fit index* (sur 5) indiquée à côté de chaque catégorie reflète de nouveau la proportion de l'identification et la note moyenne.

F2 \ F1	1200	1100	1000	900	800	700
340						o0,5
380			o0,1	o2,2	o3,2	o5
420			o0,4	o3,5	o4,7	o5
460		o0,1		o2,5	o4,5	o3,9
				<i>ɔ0,8</i>		
500			ɔ0,1	o1	o2,4	o3
				<i>ɔ2</i>	<i>ɔ0,7</i>	
540			o0,5	ɔ2,8	o0,5	o2,8
			<i>ɔ1,7</i>		<i>ɔ2,6</i>	<i>ɔ0,5</i>
580		ɔ1,8	ɔ4,1	ɔ3,7	ɔ3,1	ɔ2,9
620		ɔ3,7	ɔ4,1	ɔ4,1	ɔ3,7	ɔ3,7
660	ɔ0,5	ɔ1,9	ɔ3,1	ɔ3,7	ɔ3,4	ɔ3,4
700			ɔ0,2	ɔ0,7	ɔ3,2	ɔ3,5

Tableau 132 : Identification des voyelles synthétisées en termes de /o/ ou /ɔ/ (en italique) avec la valeur du *fit index* de l'identification, calculé sur la base de la proportion de l'identification et de la note moyenne obtenue

Le Tableau 132 montre que le contraste o/ɔ repose aussi essentiellement sur l'indice de F1 même si la frontière entre l'une et l'autre catégorie est moins bien définie et dépend de la valeur de F2. Ainsi les sons avec F1 entre 380 Hz et 460 Hz et F2 entre 700 Hz et 800 Hz sont catégorisés comme /o/. La note de qualité maximale est alors attribuée aux sons synthétisés avec F1 entre 380 et 420 Hz et F2 à 700 Hz. A l'inverse, les sons avec F1 entre 580 Hz et 660 Hz et F2 entre 700 et 1100 Hz sont catégorisés comme /ɔ/ et la note de qualité maximale est accordée aux sons produits avec F1 de 620 Hz et F2 entre 900 et 1000 Hz. Enfin le *fit index* maximal d'identification des sons en termes de /ɔ/ est de 4,1 sur 5 ce qui montre que ce son vocalique est plus difficilement identifiable que les sons catégorisés en termes de /o/ fermé (identifiés avec *fit index* maximal de 5).

Lorsque F1 se situe entre 460 et 540 Hz et le F2 entre 700 et 1000 Hz, les sons physiquement identiques peuvent renvoyer aux deux catégories /o/ et /ɔ/ et pour ces sons F2 est décisif. Par exemple lorsque F1 est de 500 Hz, les sons renvoient pour la plupart au /o/, si F2 est de 800 Hz, et au /ɔ/, si F2 est de 900 Hz. Lorsque F1 est de 540 Hz, les sons sont majoritairement interprétés en termes de /o/, si F2 est de 700 Hz, et en termes de /ɔ/, si F2 est de 800 Hz. **Au sein de ces sons 'border line', plus le F1 est élevé, plus le F2 doit être bas (et donc plus la voyelle doit avoir le caractère focal) pour identifier /o/. Ainsi, le F2 peut dans une certaine zone acoustique « compenser » le F1.**

Notons également que la plage de variation de F2 est large pour les sons identifiés en termes de /ɔ/ (elle varie entre 700 Hz et 1100 Hz) alors qu'elle est plus restreinte pour les sons catégorisés comme /o/ (le F2 varie alors pour la plupart entre 700 Hz et 900 Hz).

Nous concluons que même si le contraste o/ɔ repose essentiellement sur le premier formant, le deuxième formant a également un grand rôle perceptif.

Enfin, le Tableau 133 indique par quelles fréquences F1 et F2 se réalisent les sons catégorisés comme /ø/ ou /œ/. La valeur du *fit index* est précisée à côté de chaque catégorie et résume la proportion d'identification ainsi que la note de qualité moyenne.

F2 \ F1	2400	2000	1800	1600	1400	1200	1100	1000	900
340			ø3,2	ø5	ø4,7	ø4,8	ø4,5	ø1,5	
380		ø4,4	ø4,4	ø4,9	ø4,8	ø4,5	ø4,7	ø2,7	
420		ø2,9	ø4,7	ø4,6	ø4,8	ø4,4	ø3,6	ø2,3	
460		æ2,5	ø2 æ0,7	ø0,4 æ3	ø0,6 æ2,2	ø0,3 æ2,9	ø1 æ1,5	ø0,1 æ1,9	
500		ø1,5 æ0,4	æ3,1	ø0,6 æ2,9	æ4,2	æ3,5	æ3,5	æ2,9	
540		ø0,4 æ2,4	æ3,6	æ4,3	æ4,6	æ3,5	æ4,2		
580			æ4,5	æ4,6	æ4,9	æ4,5	æ0,7		
620			æ3,8	æ5	æ5	æ4,4			
660		æ0,1	æ4,6	æ4,8	æ5	æ2,2	æ0,2		
700			æ4,2	æ5	æ4,5				

Tableau 133 : Identification des voyelles synthétisées en termes de /ø/ ou /œ/ (en italique) avec la valeur du *fit index* de l'identification, calculé sur la base de la proportion de l'identification et de la note moyenne obtenue

Le Tableau 133 indique que le contraste $\emptyset/\text{œ}$ se base de nouveau essentiellement sur la valeur du premier formant. Alors que les sons ayant F1 entre 340 et 420 Hz et F2 entre 1000 Hz et 2000 Hz sont interprétés en termes de \emptyset /, ceux ayant F1 entre 580 Hz et 700 Hz et F2 entre 1100 Hz et 1800 Hz sont interprétés en termes de œ / (quoique si F1 est élevé et correspond à 700 Hz, F2 doit être supérieur à 1200 Hz). Enfin, les sons avec F1 entre 460 Hz et 540 Hz peuvent être interprétés en termes de \emptyset / ou œ / et la valeur de F2, lorsqu'elle se situe entre 1000 Hz et 1600 Hz n'a aucun impacte sur l'interprétation : dans cette plage et avec F1 à 460 Hz, les sons sont pour la plupart catégorisés comme \emptyset /.

Il est intéressant à noter que les sons identifiés comme \emptyset / avec un *fit index* élevé (entre 4,7 et 5 sur 5) ont F1 entre 340 Hz et 420 Hz et la plage de variation de F2 très large : entre 1100 Hz et 1800 Hz. En revanche, la plage de variation de F2 pour la perception d'un œ / authentique est moins large et varie entre 1400 et 1600 Hz. De cette façon, lorsque F2 est situé entre 1400 et 1600 Hz, il n'affecte ni la catégorisation en termes d'une ou de l'autre catégorie ni la qualité du son perçu. De cette façon nous stipulons que le contraste $\emptyset/\text{œ}$ est essentiellement basé sur F1 quoique la plage de variation de F2 pour les sons représentant la catégorie œ / soit plus restreinte que celle pour les sons représentant le \emptyset /.

Pour résumer, les résultats du test indiquent l'importance de F1 dans la perception des contrastes $\text{e}/\text{ɛ}$, $\emptyset/\text{œ}$, $\text{o}/\text{ɔ}$ mais également celle de F2 pour $\text{o}/\text{ɔ}$, et dans une moindre mesure pour $\text{e}/\text{ɛ}$. Ainsi, dans le cas de nos stimuli générés avec Klatt F1 doit être inférieur :

- à 540 Hz pour percevoir $\text{e}/$,
- à 460 Hz pour percevoir $\emptyset/$ et $\text{o}/$

Ensuite, F1 doit être supérieur à

- 540 Hz pour percevoir $\text{ɛ}/$, $\text{œ}/$ et $\text{ɔ}/$

La plage de variation de F2 se situe entre :

- 1000 Hz – 2000 Hz pour identifier $\emptyset/$ et entre 1100 Hz – 1800 Hz pour identifier $\text{œ}/$
- 700 Hz – 900 Hz pour identifier $\text{o}/$ et entre 700 Hz – 1100 Hz pour identifier $\text{ɔ}/$
- 2400 Hz – 2800 Hz pour identifier $\text{e}/$ et entre 2000 Hz – 2800 Hz pour identifier $\text{ɛ}/$

Les sons identifiés au travers des six différentes catégories avec un *fit index* de 4 et plus sur 5 se trouvent marqués dans l'espace acoustique F1/F2 de la Figure 115 à la page 294. Ceux catégorisés avec un *fit index* maximal (de 5) et donc les meilleurs représentants de la catégorie sont marqués avec une police plus grande.

La Figure 115 fait ressortir que les paires des voyelles $\text{e}/\text{ɛ}$ et $\emptyset/\text{œ}$ se distinguent essentiellement par le F1 et celles de la paire $\text{o}/\text{ɔ}$ s'opposent par les deux formants, F1 et F2.

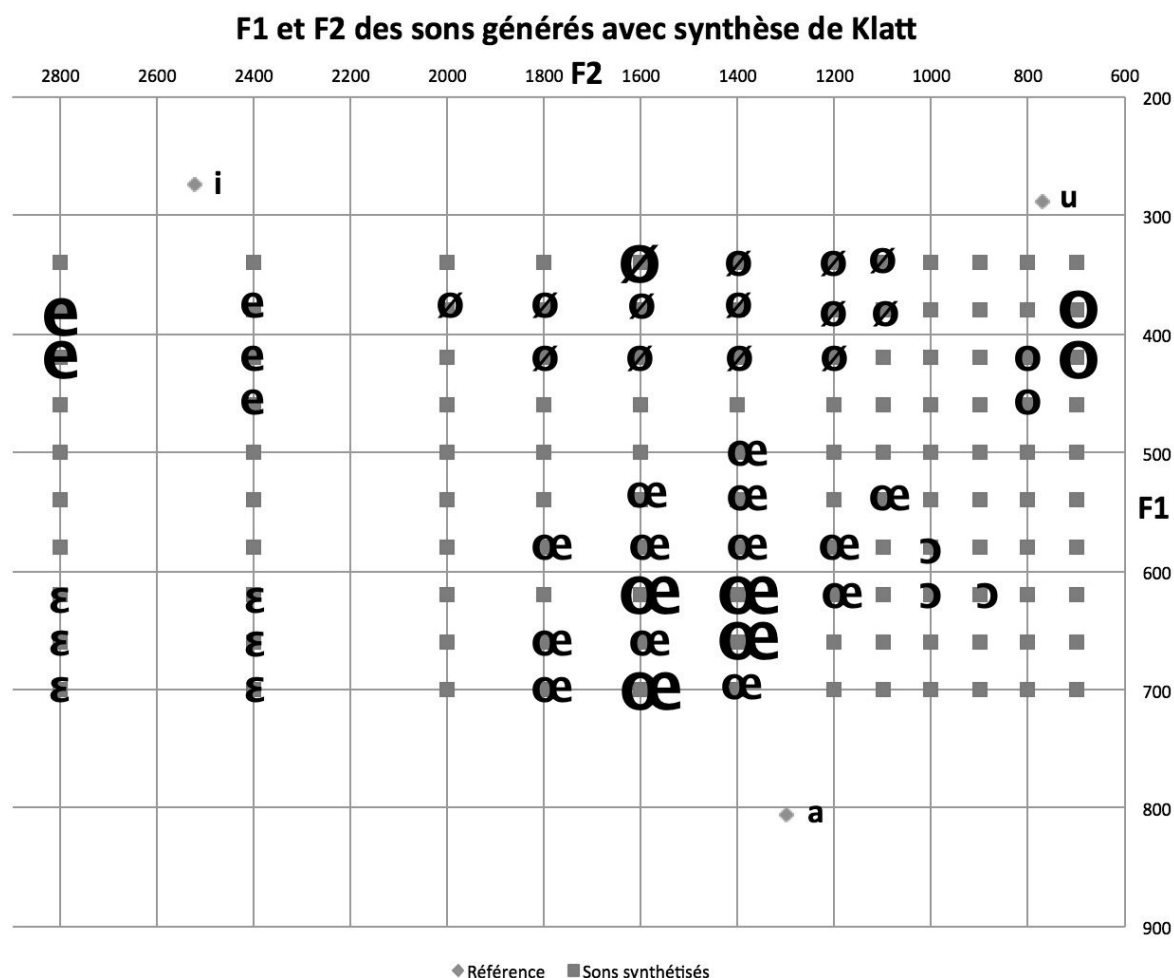


Figure 115 : Sons vocaliques synthétisés avec Klatt catégorisés comme /e, ε, ø, œ, o, ɔ/ avec un *fit index* de 4 et plus sur 5. Ceux catégorisés avec un *fit index* de 5 sont marqués en grand. Les voyelles [i, a, u] servent de point de repère

13.1.6 Perception des voyelles moyennes : Conclusion partielle

Dans l'ensemble, le résultat des cinq épreuves perceptives montre que les auditeurs français natifs testés perçoivent six voyelles distinctes /e, ε, ø, œ, o, ɔ/. Les « meilleurs » résultats sont observés dans le test 1 de discrimination (e/ε discriminées dans 99 % des réponses, o/ɔ dans 98 % des réponses et ø/œ dans 95 % de réponses) et dans le test 2 d'identification (e/ε et o/ɔ non-standard identifiés dans 100 % des réponses et ø/œ dans 94 % des réponses). Le résultat du deuxième test montre que la distribution des voyelles moyennes en français standard obéit essentiellement aux règles orthoépiques selon lesquelles /e/ n'apparaît pas en syllabe finale fermée et /œ, ɔ/ n'apparaissent pas en syllabe ouverte finale.

Le résultat des troisième et quatrième tests (identification de mots des paires minimales et identification de sens inattendu) sur le statut phonémique des voyelles moyennes révèle que le contraste e/ε peut être réalisé en syllabe finale ouverte et les contrastes ø/œ, o/ɔ en syllabe finale fermée. Cependant quoique le nombre de paires minimales théoriques attestant du statut phonémique des voyelles e/ε et o/ɔ est grand, le nombre de contextes oppositifs effectifs où les contrastes sont systématiquement perçus est selon nos résultats restreint. Les mêmes résultats montrent que les voyelles de la paire o/ɔ ont un statut phonémique plus robuste dans l'esprit des auditeurs français natifs non méridionaux (identifiées dans respectivement 93 % et 86 % des réponses) que e/ε (identifiées dans respectivement 90 % et 71 % des réponses). Ce résultat est en accord avec l'étude de (Landick, 1995)

et le questionnaire phonologique de (Martinet, 1945). Le contraste entre les voyelles *ø/œ* est également productif, notamment dans la paire *veulent/ veule* (adj.).

Dans les contextes oppositifs où les auditeurs natifs ne sont plus sensibles aux oppositions basées sur les voyelles moyennes, leur perception est majoritairement régie par la loi de position selon laquelle les voyelles moyennes de qualité mi-fermée [*e, o, ø*] apparaissent en syllabe finale ouverte et celles de qualité mi-ouverte [*ɛ, ɔ, œ*] en syllabe finale fermée. C'est notamment le cas de la voyelle <*e*> en syllabe ouverte pour laquelle le dictionnaire (www.lerobert.com) impose la prononciation particulière par le timbre mi-ouvert [*ɛ*], comme dans les mots *baie*, *taie* ou *valet*, et pour laquelle les auditeurs français testés acceptent dans les tâches 3 et 4 (selon respectivement 77 % et 64 % des réponses) en français standard la qualité mi-fermée [*e*].

Enfin, le résultat de l'identification de voyelles synthétisées avec Klatt a montré que le contraste entre les voyelles moyennes repose essentiellement sur la valeur de F1 et que F2 joue un rôle perceptif important dans la distinction de *o/ɔ* et dans l'attribution de la note de qualité aux voyelles *e/ɛ*.

13.2 Production de contrastes entre *e/ɛ, o/ɔ, ø/œ*

Pour savoir ce qu'il en est de la production des contrastes entre les voyelles moyennes chez les locuteurs français natifs, nous avons analysé :

- Les réponses de cinq Français natifs non-méridionaux à un questionnaire phonologique,
- la réalisation du contraste acoustique (en termes des valeurs formantiques moyennes) entre les voyelles *e/ɛ, o/ɔ* et *ø/œ* en contextes labial, coronal, palato-vélaire et uvulaire chez les dix Françaises natives qui constituent la "référence" dans cette thèse. Alors que ce contraste est marqué par les natives pour les voyelles produites en isolation, comme démontré en 8.3.1, nous avons voulu savoir dans quelle mesure il est réalisé dans des environnements phonétiques divers.

13.2.1 Questionnaire phonologique sur la production

Afin de déterminer si les cinq Français natifs, ayant servi dans les quatre études perceptives et dont les informations personnelles sont fournies dans le Tableau 116 à la page 277, pensent produire le contraste entre les voyelles moyennes dans certaines paires minimales, nous leur avons proposé un questionnaire phonologique.

Ce questionnaire comporte au total 30 paires minimales basées sur le contraste entre les voyelles moyennes ; les enquêtés répondaient à la question « Prononcez-vous les deux mots de chaque paire de la même façon ? » en choisissant la réponse « Toujours », « Souvent », « Rarement » ou « Jamais ». Les paires minimales proposées étaient au nombre de 14 pour les contrastes *e/ɛ* et *o/ɔ* et de deux pour le contraste *ø/œ*. Le questionnaire se trouve à la Figure 116 de la page 296.

	Toujours	Souvent	Rarement	Jamais
blé - blet				
les - lait				
créé - craie				
thé - taie				
j'aurai - j'aurais				
manger - mangeais				
votre - vautre				
pomme - paume				
heaume - homme				
jeûne - jeune				
veule (adj.) - veulent				
B - baie				
Saône - sonne				

	Toujours	Souvent	Rarement	Jamais
mes - mais				
Aude - ode				
khôl - col				
taupe - top				
hausse - os				
rauque - rock				
fée - fait				
vallée - valet				
Beauce - bosse				
pôle - Paul				
dé - dès				
allée - allait				
cote - côte				
molle - môle				
Paul - Paule				
été - étais				
tournée - tournaient				

Figure 116 : Questionnaire phonologique portant sur la façon de prononcer les deux mots de chaque paire minimale dont le contraste repose sur les voyelles moyennes

Le résultat du questionnaire fait apparaître que les cinq locuteurs français pensent ne jamais prononcer de la même façon les mots basés sur le contraste e/ɛ : 49 réponses sur 70 (70 %) ; les mots basés sur le contraste o/ɔ : 55 réponses sur 70 (79 %) ; enfin les mots basés sur le contraste ø/œ : 6 réponses sur 10 (60 %). Les résultats individuels des enquêtés se trouvent dans le Tableau 134.

13. Voyelles moyennes du français chez des Français natifs non-méridionaux

Jamais confondues	e/ɛ sur 14	o/ɔ sur 14	ø/œ sur 2
1 : Gien	5	12	2
2 : Lyon	9	13	1
3 : Paris	12	10	/
4 : Enghien les Bains	10	8	1
5 : St. Briec	12	12	2
TOTAL	49/70 (70%)	55/70 (79%)	6/10 (60%)

Tableau 134 : Nombre de paires minimales basées sur le contraste des voyelles moyennes qui ne sont jamais prononcées de la même façon selon 5 enquêtés français natifs non méridionaux

Le Tableau 135 détaille les réponses des enquêtés (le numéro de l'enquêté se trouve alors entre parenthèses) concernant la production des paires minimales basées sur le contraste *e/ɛ*.

- quatre enquêtés pensent produire de la même façon les mots de la paire minimale *mes/mais* (deux toujours et deux rarement) ;
- trois enquêtés pensent produire de la même façon les paires *j'aurai/ j'aurais* (deux toujours, un souvent), *les/ lait* (deux rarement, un souvent) ;
- deux enquêtés pensent pouvoir produire de la même manière les mots des paires *fée/ fait*, *dé/dès*, *été/ étais* (dans les trois cas un rarement, un souvent) ;
- un Français natif pense pouvoir produire souvent les mots des paires *thé/ taie et créé/ craie*, *vallée/ valet*, *allée/ allait* de la même façon et un produirait rarement les mots *ournée/ tournaient* avec le même timbre.

En revanche, les cinq enquêtés pensent ne jamais produire de la même manière les trois paires minimales suivantes basées sur le contraste *e/ɛ* : *blé/ blet*, *manger/ mangeais*, *B/baie*.

Confusion e/ɛ	Toujours	Souvent	Rarement
Mes/mais	✓ (1)	✓ (4)	✓ (2, 5)
J'aurai/ j'aurais	✓ (2, 3)	✓ (1)	
Les/lait		✓ (4)	✓ (1, 5)
Fée/ fait		✓ (1)	✓ (2)
Dé/ dès		✓ (1)	✓ (2)
Été/ étais		✓ (1)	✓ (2)
Valée/ valet		✓ (1)	
Allée/ allait		✓ (1)	
Thé/ taie		✓ (4)	
Créé/ craie			✓ (4)
Tournée/ tournaient			✓ (1)

Tableau 135 : La même réalisation des mots de 14 paires minimales dont le contraste repose sur les voyelles moyennes *e/ɛ*. Questionnaire rempli par 5 Français natifs non-méridionaux ; le no de l'auditeur ayant choisi la réponse indiquée se trouve entre parenthèses

Le Tableau 136 de la page 298 renseigne ensuite sur la production par cinq enquêtés des paires minimales basées sur le contraste des voyelles moyennes *o/ɔ*.

- trois enquêtés pensent pouvoir rarement prononcer les mots *Paul/ Paule* de la même façon ;

- deux enquêtés pensent produire de la même façon les mots des paires *Aude/ ode* (toujours), *pôle/ Paul* (les deux rarement) et *vautre/ votre* (un toujours, un rarement) ;
- un enquêté pense souvent prononcer de la même manière *khôl/ col*, *hausse/ os*, *Beauce/ bosse* et rarement *môle/ molle*, *taupe/ top* et *heaume/ homme*.

Tous les cinq pensent en revanche faire la différence entre les mots des quatre paires minimales suivantes : *pomme/ paume*, *Saône/ sonne*, *rauque/ rock* et *cote/ côte*.

Confusion o/ɔ	Toujours	Souvent	Rarement
Pôle/ Paul			✓ (1, 4)
Aude/ ode	✓ (3, 4)		
Paul/ Paule			✓ (1, 5, 2)
Vautre/ votre	✓ (3)		✓ (4)
Khôl/ col		✓ (3)	
Hausse/ os		✓ (5)	
Beauce/ bosse		✓ (4)	
Môle/ molle			✓ (4)
Taupe/ top			✓ (4)
Heaume/ homme			✓ (3)

Tableau 136 : Réalisation identique des mots de 14 paires minimales dont le contraste repose sur les voyelles moyennes o/ɔ. Questionnaire rempli par 5 Français natifs non-méridionaux : le no de l'auditeur ayant choisi la réponse indiquée se trouve entre parenthèses

Enfin, le Tableau 137 renseigne sur la production des mots dont le contraste repose sur les voyelles moyennes ø/œ. Nous remarquons que les deux paires minimales *jeûne/ jeune* et *veule/ veulent* pourraient être produites par deux locuteurs natifs de la même façon (toujours ou rarement).

Confusion ø/œ	Toujours	Souvent	Rarement
Jeûne/ jeune	✓ (3)		✓ (2)
Veule/ veulent	✓ (3)		✓ (4)

Tableau 137 : Réalisation identique des mots de 2 paires minimales dont le contraste repose sur les voyelles moyennes ø/œ. Questionnaire rempli par 5 Français natifs non-méridionaux : le no de l'auditeur ayant choisi la réponse indiquée se trouve entre parenthèses

Le questionnaire phonologique montre que le contraste le plus productif selon les cinq enquêtés serait celui de o/ɔ (car ces voyelles ne seraient jamais prononcées de la même façon dans 79 % des réponses), puis e/ɛ (car jamais prononcées de la même façon dans 70 % des réponses) et enfin le contraste le plus fragile serait celui du ø/œ (car jamais prononcées de la même façon dans seulement 60 % des réponses). Ces résultats sont en concordance avec le questionnaire phonologique de Martinet (1945, p. 130) qui montre que la distinction o/ɔ repose sur une différence phonétique claire chez les locuteurs qui viennent d'une région autre que le Midi et qui stipule, concernant les phonèmes /ø, œ/, que « *Il y a donc en français un embryon de distinction phonologique entre les deux timbres de Œ.* ».

13.2.2 Contrastes en contextes p, t, k, R entre e/ε, ø/œ (F1 - F3) et o/ɔ (F1, F2)

Afin de déterminer l'évolution des valeurs formantiques des voyelles moyennes du français produits par les dix Françaises natives en isolation et en syllabe finale CVC où C représente l'une des consonnes [p, t, k, ʁ], nous avons effectué une analyse des formants de chacune des six voyelles [e, ε, o, ɔ, ø, œ], puis comparé leurs valeurs au sein de chaque couple e/ε, o/ɔ, ø/œ. Le corpus, l'analyse et les informations sur les locutrices sont décrits dans le chapitre 6.

13.2.2.1 Contraste e/ε

La Figure 59 illustre les valeurs formantiques moyennes du [e] mi-fermé (marqué en rouge par un rond) et [ε] mi-ouvert (marqué en vert par une croix), calculées à partir de quatre productions par dix Françaises natives non-méridionales en isolation (traits horizontaux) et dans des logatomes CVCVCV où C est la consonne [p, t, k, ʁ].

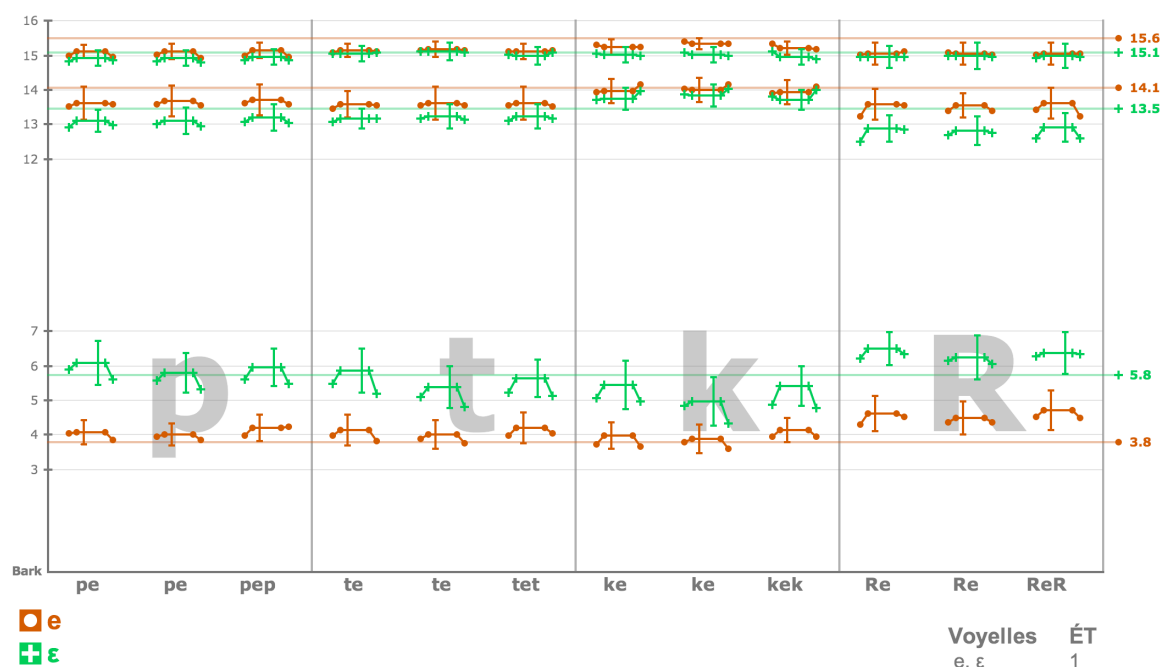


Figure 117 : Contraste acoustique entre les voyelles e/ε produites en contexte nul (traits horizontaux) et dans des logatomes CVCVCV où C = [p, t, k, ʁ]. Valeurs de F1, F2 et F3 (en Bark) sont calculées à partir de 4 répétitions par 10 Françaises natives et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1

Selon la Figure 59 le contraste entre les voyelles de la paire e/ε en syllabe finale CVC où C correspond aux consonnes [p, t, k, ʁ] est essentiellement assuré par la valeur du premier formant, puis par celle du deuxième formant. Le [e] mi-fermé se réalise alors dans tous les contextes avec F1 plus bas et F2 plus élevé que le [ε] mi-ouvert. Les valeurs moyennes des formants sont ensuite récapitulées dans le Tableau 138 de la page 300. La première colonne renseigne sur les valeurs en Hertz, la deuxième sur les valeurs transformées en Bark et enfin la troisième colonne indique la différence (en pourcentage) entre les formants de chaque couple.

La comparaison des valeurs formantiques des deux voyelles indique qu'il existe dans tous les contextes une différence supérieure à 20 % au niveau de leur premier formant et inférieure à 5 % au niveau de leur deuxième et troisième formant

		F1 (Hz)	F2 (Hz)	F3 (Hz)	F1 (Bark)	F2 (Bark)	F3 (Bark)	Diff F1 %	Diff F2 %	Diff F3 %
o	e	400	2503	3361	9,4	23	25,6			
	ε	624	2244	3065	12,2	22,1	24,8	29,8	-4	-3,1
p	e	438	2324	3059	9,9	22,4	24,8			
	ε	647	2120	2941	12,5	21,6	24,4	26,3	-3,6	-1,6
t	e	440	2288	3039	10	22,3	24,7			
	ε	608	2125	2965	12,1	21,7	24,5	21	-2,7	-0,8
k	e	431	2420	3110	9,8	22,8	24,9			
	ε	580	2312	2930	11,8	22,4	24,4	20,4	-1,8	-2
R	e	499	2279	3003	10,8	22,3	24,6			
	ε	700	2021	2965	13	21,2	24,5	20,4	-4,9	-0,4

Tableau 138 : Les valeurs formantiques F1, F2 et F3 des voyelles [e] et [ε] produits par 10 Françaises natives (4 répétitions) en isolation et en syllabe finale pVp, tVt, kVk, RVR (valeur prise à la moitié de la durée du segment), en Hertz (1^e colonne), en Bark (2^e colonne) et la différence en pourcentage (3^e colonne)

En suivant la *différence juste perceptible* (JND) > 5 %, telle qu'elle a été définie par Flanagan (1955), nous considérons que seul le F1 participe à la différenciation du contraste e/ε et du point de vue acoustico-perceptif et le contraste peut alors être décrit comme suit :

$$F1 [e] < F1 [\epsilon]$$

Cette définition est en accord avec les résultats du test d'identification des voyelles synthétisées avec Klatt (section 13.1.5).

Le contraste entre les voyelles e/ε réalisé par une Française native en isolation et en syllabe finale CVC où C est une des consonnes [p, t, k, ʀ] est illustré dans le spectrogramme de la Figure 118.

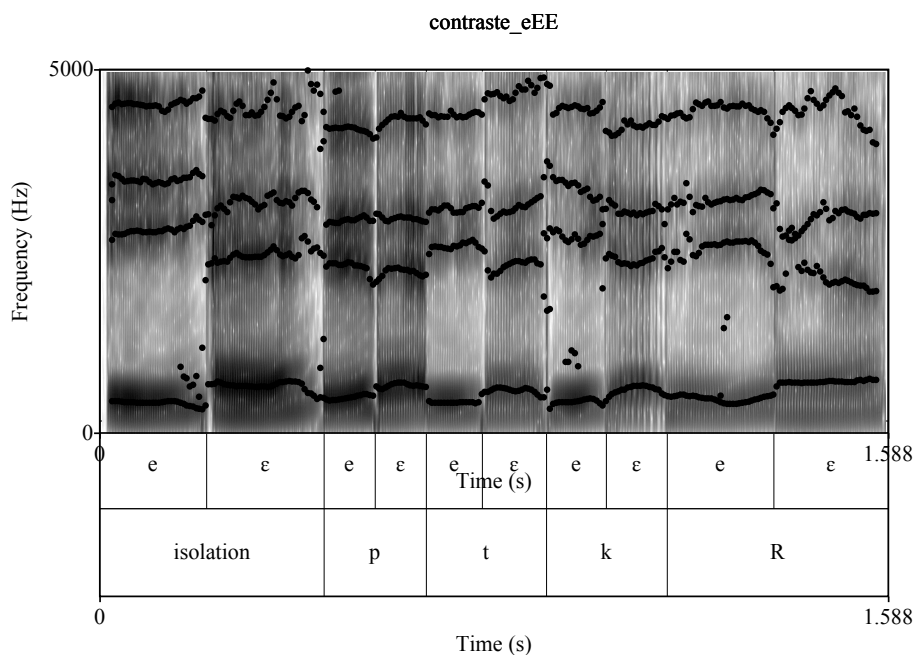


Figure 118 : Spectrogramme des voyelles [e] et [ε] produites par une Française native en isolation et en contextes symétriques p, t, k, R (syllabe finale)

Nous remarquons chez une Française native (Figure 118) que le contraste entre les voyelles e/ɛ se réalise systématiquement dans tous les environnements phonétiques avec F1 de [e] inférieur à F1 de [ɛ]. Les valeurs relevées à la moitié de la durée de la voyelle sont précisées dans le Tableau 139.

		F1 (Hz)	F2 (Hz)	F3 (Hz)
o	e	418	2770	3445
	ɛ	644	2418	3225
p	e	498	2343	2932
	ɛ	698	2256	2979
t	e	419	2579	3038
	ɛ	604	2349	3115
k	e	471	2656	3362
	ɛ	656	2343	3020
R	e	477	2593	3155
	ɛ	719	2081	3050

Tableau 139 : Les valeurs formantiques F1, F2 et F3 des voyelles [e] et [ɛ] produits par une Française native en isolation et en syllabe finale pVp, tVt, kVk, RVR (valeur prise à la moitié de la durée du segment)

Il est à noter que selon le contexte phonétique (nul ou [p, t, k, ʁ]), le contraste entre e/ɛ peut être marqué par la Française native au niveau du deuxième et/ou troisième formant avec F1 qui reste l'indice acoustique principal dans la distinction des deux qualités.

13.2.2.2 Contraste o/ɔ

La Figure 119 illustre ensuite le contraste acoustique entre [o] mi-fermé (marqué par un rond) et [ɔ] mi-ouvert (marqué par une croix). Les valeurs formantiques moyennes sont calculées à partir des quatre répliques de chacune des voyelles par dix Françaises natives non-méridionales en isolation (traits horizontaux) et dans des logatomes CVCVCVC où C est la consonne [p, t, k, ʁ].

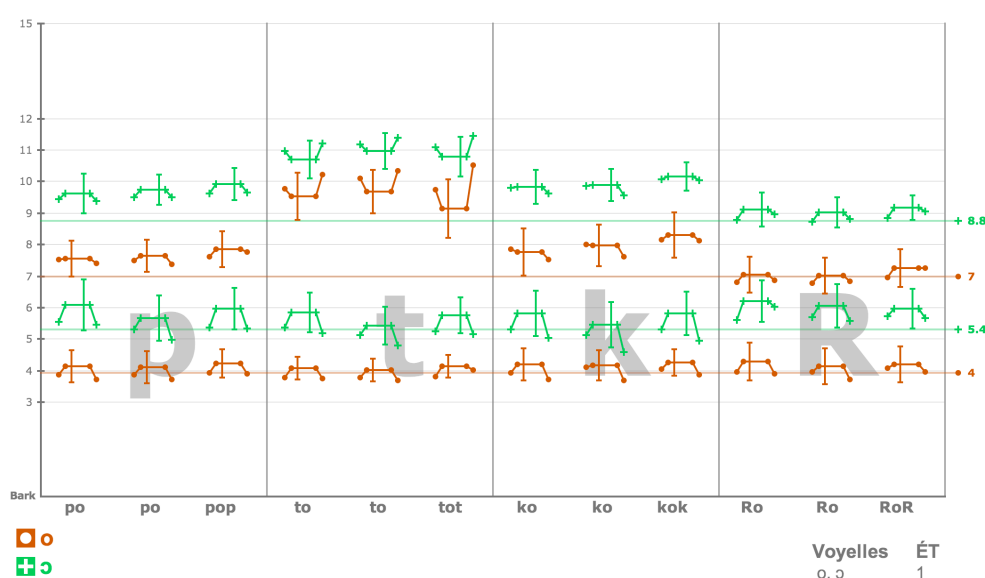


Figure 119 : Contraste acoustique entre les voyelles o/ɔ produites en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ]. Valeurs de F1, F2 (en Bark) sont calculées à partir de 4 répétitions par 10 Françaises natives et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1

Selon la Figure 119, le contraste acoustique des voyelles o/ɔ repose sur la valeur des premier et deuxième formants. Ainsi, les formants F1 et F2 du [o] mi-fermé sont plus bas que ceux du [ɔ] mi-ouvert. Le Tableau 140 fournit les moyennes pour chaque contexte séparément : en Hertz (colonne 1) et en Bark (colonne 2). La différence au niveau de chaque formant entre [o] et [ɔ], exprimée en pourcentage, est indiquée en colonne 3 et enfin la distance F2-F1 (en Hertz) de chaque voyelle permettant de rendre compte du caractère focal se trouve en colonne 4.

		F1 (Hz)	F2 (Hz)	F1 (Bark)	F2 (Bark)	Diff F1 %	Diff F2 %	F2-F1 (Hz)
0	o	417	790	4	7			373
	ɔ	577	1069	5,4	8,8	25,9	20,5	492
p	o	443	907	4,2	7,8			464
	ɔ	649	1261	6	9,9	30	21,2	612
t	o	433	1119	4,1	9,1			686
	ɔ	625	1440	5,8	10,8	29,3	15,7	815
k	o	448	979	4,3	8,3			531
	ɔ	632	1304	5,8	10,1	25,9	17,8	672
R	o	440	819	4,2	7,2			379
	ɔ	651	1122	6	9,2	30	21,7	471

Tableau 140 : Les valeurs formantiques F1 et F2 des voyelles [o] et [ɔ] produits par 10 Françaises natives (4 répétitions) en isolation et en syllabe finale pVp, tVt, kVk, RVR (valeur prise à la moitié de la durée du segment), en Hertz (1^e colonne), en Bark (2^e colonne), la différence en pourcentage (3^e colonne) et la distance F2-F1 (4^e colonne)

Le Tableau 140 montre une différence des deux premiers formants du [ɔ] par rapport à ceux du [o] : elle est supérieure à 25 % pour le F1 et à 15 % pour le F2, selon le contexte étudié. Ensuite, la distance entre F2 et F1 est dans tous les contextes plus réduite pour [o] que pour [ɔ]. Ainsi, du point de vue acoustico-perceptif, nous définissons le contraste o/ɔ de la façon suivante :

$$\begin{aligned} F1 [o] &< F1 [\varepsilon] \\ F2 [o] &< F2 [\varepsilon] \\ (F1F2) [o] &< (F1F2) [\varepsilon] \end{aligned}$$

Cette définition correspond aux résultats du test d'identification des voyelles synthétisées avec Klatt (section 13.1.5).

Le contraste entre les voyelles o/ɔ réalisé par une Française native en isolation et en syllabe finale CVC avec C qui représente la consonne [p, t, k, ʁ] est illustré dans le spectrogramme de la Figure 120.

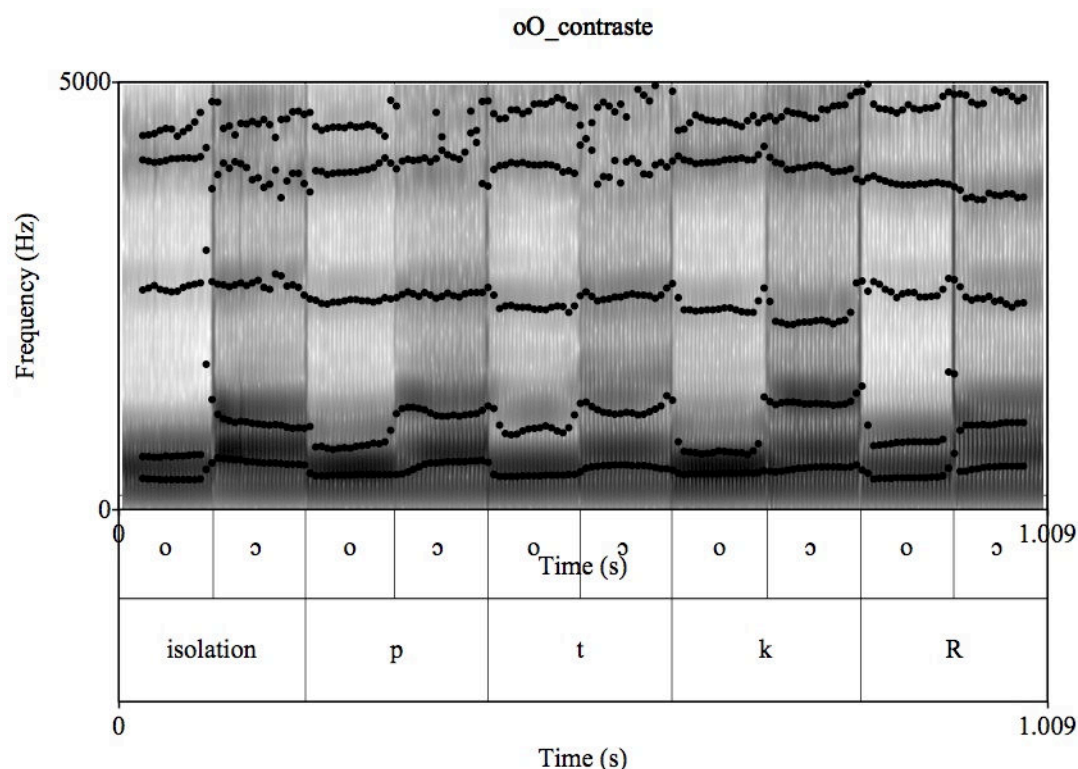


Figure 120 : Spectrogramme des voyelles [o] et [ɔ] produites par une Française native en isolation et en contextes symétriques p, t, k, R (syllabe finale)

Nous remarquons chez cette Française native (Figure 120) que le contraste o/ɔ repose acoustiquement sur les valeurs de F1 et F2 qui sont plus basses pour [o] que pour [ɔ] (voir le Tableau 141).

		F1 (Hz)	F2 (Hz)
o	o	350	628
	ɔ	554	997
p	o	405	728
	ɔ	546	1092
t	o	395	931
	ɔ	520	1133
k	o	422	683
	ɔ	498	1240
R	o	372	792
	ɔ	502	998

Tableau 141 : Les valeurs formantiques F1 et F2 (en Hertz) des voyelles [o] et [ɔ] produits par une Française native en isolation et en syllabe finale pVp, tVt, kVk, RVR (valeur prise à la moitié de la durée du segment)

Le Tableau 141 indique que les valeurs relevées pour la locutrice native vont dans le sens de la tendance générale, à savoir une valeur plus basse de F1 et F2 pour [o] par rapport au [ɔ] et ceci dans tous les contextes étudiés.

13.2.2.3 Contraste ø/œ

Enfin, la Figure 121 indique la distance acoustique (en termes des formants) entre [ø] mi-fermé (marqué par un rond) et [œ] mi-ouvert (marqué par une croix) dont les valeurs formantiques moyennes ont été calculées à partir de quatre répétitions de chacune des voyelles par dix Françaises natives non-méridionales en isolation (traits horizontaux) et dans des logatomes CVCVCVC où C correspond à [p, t, k, ʁ].

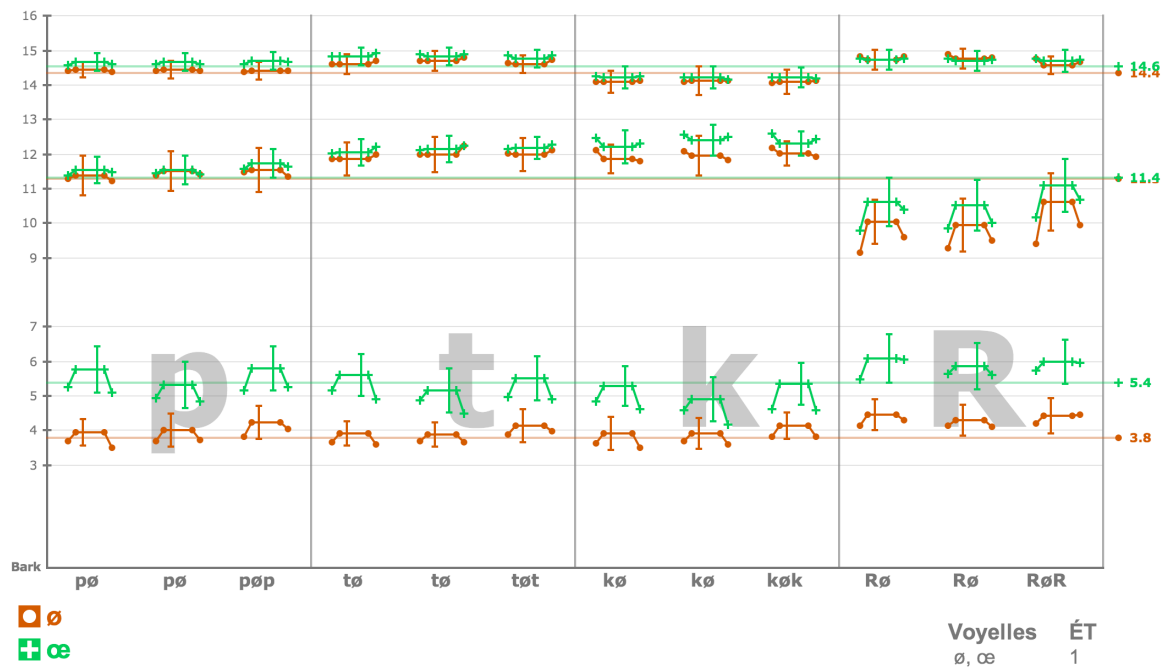


Figure 121 : Contraste acoustique entre les voyelles ø/œ produites en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ]. Valeurs de F1, F2 et F3 (en Bark) sont calculées à partir de 4 répétitions par 10 Françaises natives et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1

Comme on peut le voir à la Figure 121, le contraste ø/œ est essentiellement assuré par la valeur du premier formant et selon le contexte phonétique également par les valeurs du deuxième et dans une moindre mesure du troisième formant. Le Tableau 142 indique les valeurs moyennes des trois premiers formants en Hertz (colonne 1) et en Bark (colonne 2). La différence entre les formants du couple est donnée en pourcentage en colonne 3.

Le Tableau 142 indique que la différence entre ø/œ susceptible d'être perçue concerne seulement le premier formant car elle est supérieure à 22 %. La différence entre les valeurs du deuxième de même que du troisième formants étant inférieure à 5 % n'est pas prise en compte dans la description acoustico-perceptive du contraste ø/œ que nous définissons comme suit :

F1 [ø] < F1 [œ]

Il est à noter que cette définition correspond au résultat du test d'identification des voyelles synthétisées avec Klatt (section 13.1.5).

Le spectrogramme de la Figure 122 illustre ensuite la réalisation du contraste ø/œ par une Françaises native. Nous remarquons que le contraste ø/œ repose essentiellement sur la valeur de F1 qui est plus basse dans le cas de [ø].

13. Voyelles moyennes du français chez des Français natifs non-méridionaux

		F1 (Hz)	F2 (Hz)	F3 (Hz)	F1 (Bark)	F2 (Bark)	F3 (Bark)	Diff F1 %	Diff F2 %	Diff F3 %
0	ø	399	1569	2642	3,8	11,3	14,4			
	œ	584	1580	2749	5,4	11,4	14,6	29,6	0,9	1,4
p	ø	442	1623	2645	4,2	11,5	14,4			
	œ	629	1666	2790	5,8	11,7	14,7	27,6	1,7	2
t	ø	432	1734	2732	4,1	12	14,6			
	œ	591	1786	2826	5,5	12,2	14,7	25,5	1,6	0,7
k	ø	432	1740	2491	4,1	12	14,1			
	œ	572	1823	2538	5,3	12,3	14,2	22,6	2,4	0,7
R	ø	464	1401	2721	4,4	10,6	14,6			
	œ	652	1505	2808	6	11,1	14,7	26,7	4,5	0,7

Tableau 142 : Les valeurs formantiques F1, F2 et F3 des voyelles [ø] et [œ] produites par 10 Françaises natives (4 répétitions) en isolation et en syllabe finale pVp, tVt, kVk, RVR (valeur prise à la moitié de la durée du segment), en Hertz (1^e colonne), en Bark (2^e colonne) et la différence en pourcentage (3^e colonne)

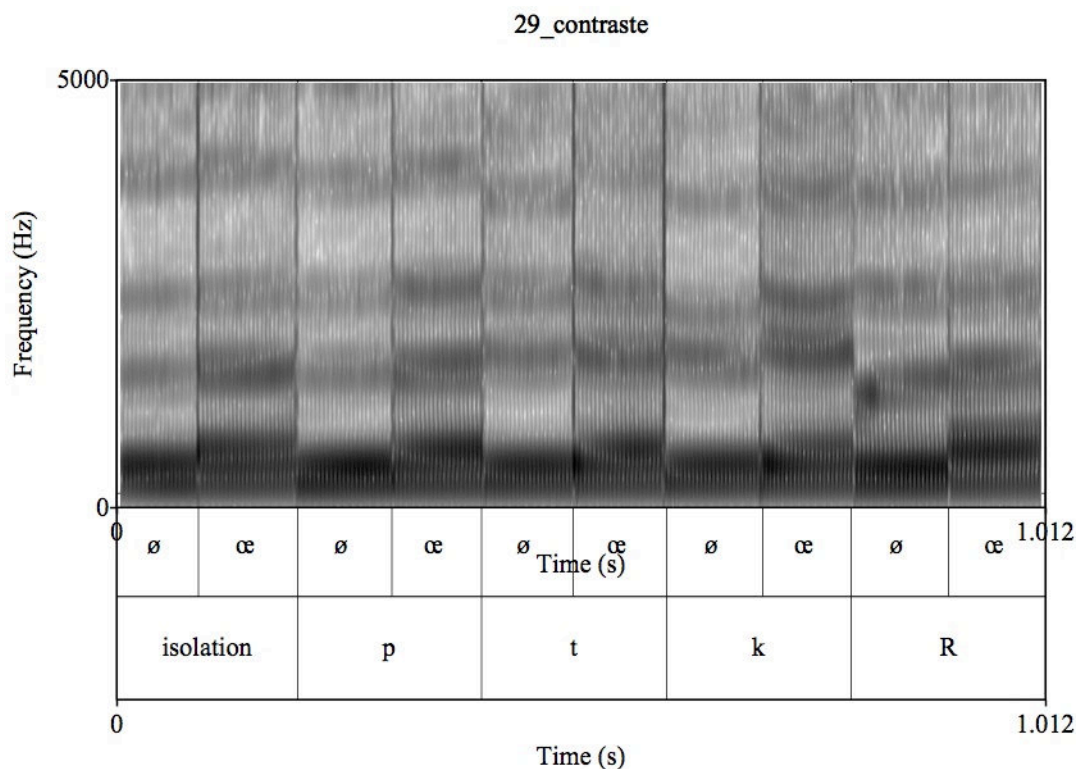


Figure 122 : Spectrogramme des voyelles [ø] et [œ] produites par une Française native en isolation et en contextes symétriques p, t, k, R (syllabe finale)

Le Tableau 143 renseigne sur les valeurs précises de F1, F2 et F3 de [ø] et [œ] produits en isolation et en syllabe finale CVC où C correspond à une des consonnes [p, t, k, ʁ] par la Française native.

		F1 (Hz)	F2 (Hz)	F3 (Hz)
0	ø	393	1373	2353
	œ	493	1546	2554
p	ø	384	1303	2141
	œ	542	1657	2612
t	ø	394	1703	2190
	œ	454	1748	2643
k	ø	406	1604	2052
	œ	454	1676	2211
R	ø	387	1282	2473
	œ	606	1601	2574

Tableau 143 : Valeurs moyennes de F1, F2 et F3 (en Hz) de [ø] et [œ] produits par une Française native en isolation et en syllabe finale pVp, tVt, kVk, RVR (valeur prise à la moitié de la durée du segment)

Ainsi le contraste entre ø/œ chez la Française native est systématiquement basé sur une différence de F1, alors que la différence au niveau de F2 et F3 peut être, selon le contexte, minime (voir la valeur du F2 en cotexte [t, k] ou la valeur du F3 en contexte [ʁ] dans le Tableau 143).

13.2.3 Production des voyelles moyennes : Conclusion partielle

Selon les réponses de cinq Français natifs au questionnaire phonologique, le contraste o/ɔ serait le plus « robuste » (à priori réalisé dans 77 % de paires minimales présentées), suivi du contraste e/ɛ (réalisé dans 70 % des paires minimales) et enfin le contraste ø/œ serait le moins actif car à priori réalisé dans seulement 60 % de paires minimales présentées.

L'analyse et la comparaison des formants des six voyelles moyennes répétées quatre fois en syllabes finales des logatomes par dix Françaises natives montre que du point de vue acoustico-perceptif, les contrastes peuvent être définis comme suit :

e/ɛ	F1 [e] < F1 [ɛ]
o/ɔ	F1 [o] < F1 [ɔ], F2 [o] < F2 [ɔ], (F1F2) [o] < (F1F2) [ɔ]
ø/œ	F1 [ø] < F1 [œ]

Ainsi, les Français non-méridionaux produisent six qualités vocaliques distinctes correspondant aux voyelles moyennes du français et ils pensent réaliser les contrastes dans la majorité des paires minimales présentées.

13.3 Voyelles moyennes chez des Français natifs : Conclusion partielle

Le chapitre 13 a permis d'avancer que les Français natifs non méridionaux perçoivent et produisent six qualités vocaliques différentes correspondantes aux voyelles moyennes [e, ε, o, ɔ, ø, œ] dont la distribution est loin d'être aléatoire. Alors que les contrastes e/ε et ø/œ reposent sur l'indice acoustique de F1, le contraste o/ɔ (le plus robuste selon les réponses) repose sur les indices acoustiques F1 et F2.

Nous avons ainsi montré que dans l'acquisition des voyelles moyennes du français, il faut veiller à :

1. Réaliser les [e] et [ø] avec F1 (corrélé essentiellement à l'aperture) plus bas que celui des [ε] et [œ], puis le [o] avec F1 et F2 (corrélé à l'antériorité-postériorité et arrondissement) plus bas (et une distance F2-F1 plus petite) que ceux du [ɔ]. Puis, pour atteindre l'authenticité, réaliser le [e] acoustiquement diffus et le [ε] acoustiquement plus compacte.

2. Suivre en premier lieu les règles orthoépiques qui sont à la base de la distribution semi-complémentaire des voyelles moyennes du français. Selon nos résultats, ces règles régissent pratiquement toujours la perception des natifs, [e] n'apparaissant jamais en syllabe finale fermée et [ɔ, œ] n'apparaissant pas en syllabe finale ouverte en français standard. Il est donc impératif de réaliser la qualité [ε] en syllabe CVC et la qualité [o] et [ø] en syllabe CV.

3. Réaliser le contraste e/ε en syllabe finale ouverte et ø/œ de même que o/ɔ en syllabe finale fermée dans quelques contextes oppositifs dont le nombre est restreint (règles particulières). Le résultat de notre étude montre que le contraste doit être systématiquement réalisé dans les quatre paires suivantes où l'inversion du timbre choque tous les auditeurs français testés :

- a. homme/ heaume
- b. top/ taupe
- c. pomme/ paume
- d. veulent/ veule (adj.)

Bien que le statut phonémique des voyelles moyennes n'ait pas été testé dans toutes les paires minimales pour lesquelles les dictionnaires²⁹ prescrivent deux timbres différents et que notre liste ne soit pas complète, nous avons mis en évidence par ce résultat la tendance vers l'affaiblissement de ces contrastes en français actuel.

4. Dans les autres contextes où les contrastes ne sont plus actifs et afin de ne pas imposer des règles obsolètes, il est possible de suivre la loi de position qui est une tendance de plus en plus répandue en français actuel tout en respectant les contraintes de la distribution des timbres en fonction du contexte phonémique, résumées par Léon (1976). Selon nos résultats, cette tendance est la plus productive pour [e] qui est largement accepté en syllabe finale ouverte des mots pour lesquels les dictionnaires prescrivent la qualité mi-ouverte. Ces résultats sont en concordance avec le modèle didactique de Wioland (2005) qui recommande la fermeture des voyelles moyennes en syllabe finale CV.

²⁹ Comme le dictionnaire en ligne Le Grand Robert : www.lerobert.com

14. Voyelles moyennes du français chez des futures enseignantes de FLE de LM tchèque

Les limites de mon langage signifient les limites de mon propre monde.

Ludwig Wittgenstein

Résumé : Le **chapitre quatorze** est une étude qualitative de la perception et production des voyelles moyennes par nos futures enseignantes. Le résultat du test de discrimination des voyelles moyennes en contexte [p, t, k, ʁ] met en évidence la difficulté à percevoir le contraste entre e/ɛ, ø/œ et o/ɔ qui varie selon les auditrices. La discrimination globale de ø/œ par les futures enseignantes est de 21%, celle de e/ɛ est de 24% et enfin celle de o/ɔ de 33%. Une auditrice discrimine cependant e/ɛ dans 14 suites sur 16 et o/ɔ dans 10 suites sur 16 ce qui prouve que l'acquisition des deux contrastes en perception est possible pour les tchécophones.

En production, nous avons analysé dix répétitions de chacune des voyelles moyennes en isolation et en syllabe finale CVC où C correspond à la consonne [p, t, k, ʁ] pour chaque future enseignante, séparément. Le timbre des voyelles isolées est ensuite vérifié par des auditeurs français natifs. Le résultat de l'étude des voyelles isolées montre que [ø] et [ɛ] sont produits par la plupart des futures enseignantes avec authenticité, [o] et [e] sont produits avec authenticité par une future enseignante, [œ] est maîtrisé sans authenticité par une enseignante et [ɔ] n'est pas maîtrisé. Le manque d'authenticité du [œ] et l'absence de maîtrise du [ɔ] isolés peuvent s'expliquer par les règles orthoépiques du français qui interdisent la distribution de ces qualités en syllabe finale CV, et donc en isolation, et pour cette raison ne doivent pas être considérés comme de l'échec dans l'apprentissage phonétique du français. Le résultat de l'analyse des formants des voyelles produites en syllabe finale CVC montre que les futures enseignantes peuvent posséder (selon la locutrice, le couple des voyelles et le contexte) deux catégories mentales pour les voyelles en opposition. La réalisation des contrastes est facilitée dans les contextes où les voyelles de chaque couple sont acoustiquement et/ou perceptivement *similaires* à deux voyelles tchèques différentes. Cependant les différences acoustiques des voyelles en contraste produites par les futures enseignantes risquent de ne pas être phonétiquement pertinentes car elles sont soit moindres par rapport à la "référence", soit concernent un indice acoustique différent de celui de la "référence" (par exemple F2 à la place de F1 pour opposer ø/œ, ou seulement F1 à la place de F1 et F2 pour opposer o/ɔ). Ainsi, quoique la visualisation des formants montre que l'apprentissage phonétique des contrastes est en cours, au moins dans certains contextes et chez certaines futures enseignantes, les oppositions doivent être acoustiquement renforcées pour être perçues.

14.1 Perception: tâche de la singularité

De nombreux chercheurs s'accordent sur le fait que la performance en perception précède celle en production dans l'apprentissage d'une LE (Landeracy, 1976; Lebel, 1969; Magnen *et al.*, 2005). L'avantage d'effectuer un test de perception en même temps que d'analyser acoustiquement la production chez la même future enseignante réside ainsi dans la possibilité d'étudier le rapport direct entre perception et production (Flege, 2003).

Les résultats du chapitre 12 ont révélé la difficulté des futures enseignantes à identifier les voyelles moyennes du français. Mais qu'en est-il de leur discrimination ? Afin de répondre à la question, les futures enseignantes tchèques ont accompli la tâche de la singularité décrite dans le Tableau 117, auparavant effectuée par cinq Français natifs avec succès (taux de discrimination entre 94 % et 99 % selon la paire des voyelles moyennes).

Les résultats de la tâche de la singularité, détaillés auditrice par auditrice, sont exposés dans le Tableau 144.

	e/ɛ sur 16	o/ɔ sur 16	ø/œ sur 16
T1	6	2	4
T2	2	4	4
T3	2	4	3
T4	5	9	6
T5	1	4	2
T6	1	5	6
T7	3	6	2
T8	2	4	2
T9	2	5	4
T10	14	10	1
TOTAL	38/160 (24%)	53/80 (33%)	34/80 (21%)

Tableau 144 : Résultat de la tâche de la singularité sur la discrimination des voyelles moyennes dans des suites de 4 monosyllabes, effectuée par 10 auditrices tchèques, futures enseignantes de FLE. Les meilleurs résultats de discrimination des deux timbres sont en gras

Le résultat du Tableau 144 montre que les futures enseignantes tchèques de FLE confondent largement les voyelles moyennes. La plupart discriminent plus souvent les voyelles o/ɔ (discriminées au total dans 33 % des réponses) que e/ɛ (discriminées dans 24 % des réponses) et que ø/œ (discriminées dans 21 % des réponses). Ce résultat de la plupart des futures enseignantes va dans le sens de la robustesse des trois contrastes chez les natifs en français actuel, avec le contraste entre o/ɔ le plus robuste et celui entre ø/œ le plus fragile. Remarquons cependant que l'auditrice T1 discrimine le plus souvent le contraste entre e/ɛ, puis ø/œ et enfin o/ɔ. De même l'auditrice T10 discrimine plus souvent les voyelles e/ɛ (contraste perçu dans 14 suites sur 16, donc le meilleur score de discrimination) que o/ɔ (discriminées dans 10 suites sur 16 qui est de nouveau le meilleur score de discrimination de ce contraste). Le contraste o/ɔ est ensuite le mieux discriminé par l'auditrice T4 (dans 9 suites sur 16). Enfin, le contraste ø/œ est le plus souvent perçu par les auditrices T4 et T6 qui discriminent les deux qualités vocaliques dans 6 suites sur 16.

Si la perception permet de prédire la performance en production, les contrastes e/ɛ et o/ɔ devraient être réalisés par la Tchèque T10 et le contraste o/ɔ éventuellement aussi par la Tchèque T4. En revanche, si l'on se fie aux résultats en perception, la réalisation du contraste ø/œ pas prévue car toutes les dix futures enseignantes confondent largement les voyelles de ce contraste. Néanmoins, puisque les auditrices T4 et T6 présentent une sensibilité plus élevée à ce contraste que les autres futures enseignantes de FLE, elles pourraient être plus performantes dans la réalisation de ce dernier.

14.2 Production: analyse formantique

L'on sait que durant l'apprentissage phonétique d'une LE, les apprenants passent par des phases d'approximation des catégories visées (Flege, 1980). Dans ces étapes, les apprenants produisent

souvent une large variété de sons pour une catégorie visée et les productions vacillantes témoignent du lancement du processus de l'apprentissage (Nemser, 1993). En ce sens, les productions des apprenants ne devraient pas être classées comme soit correctes, soit incorrectes : dans l'idéal les enseignants devraient apprécier le passage d'une mauvaise production vers une production moins mauvaise par exemple (Dickerson, 1975). Or le progrès phonétique par petits pas n'est pas facile à percevoir par l'oreille (Flege, 1980). Ainsi afin de déterminer si l'apprentissage phonétique des voyelles moyennes est en cours, une étude acoustique plus fine est nécessaire.

Dans ce but, les dix futures enseignantes de FLE de LM tchèque ont enregistré six répétitions supplémentaires (en plus des quatre déjà analysées dans le chapitre 12) des voyelles moyennes du français en isolation et en contextes [p, t, k, ʁ]. Dans le cas des voyelles insérées dans des logatomes trisyllabiques, seule la voyelle de la syllabe finale (position prosodique forte) a été étudiée et comparée sur le plan des formants avec la "référence". Le corpus ainsi que le calcul des valeurs formantiques moyennes sont récapitulés dans le Tableau 145.

Corpus voyelles isolées	6 voyelles moyennes du français, insérées dans des phrases cadre, exemple : « Père, il a dit <è> comme dans père. », répétées 10 fois
Corpus voyelles en contexte	6 voyelles moyennes du français, insérées dans des logatomes pVpVpVp, tVtVtVt, kVkVkVk, RVRVRVR d'une phrase cadre, exemple : « Le mot tototote peut bien coller. », répétées 10 fois
Locutrices	10 Tchèques natives de la région de Bohême, futures enseignantes de FLE (niveau du français avancé)
Calcul de la moyenne formantique pour chaque locutrice	Voyelles isolées : à partir de 3 valeurs, relevées à un tiers, à la moitié, et à 2 tiers de la durée de chaque voyelle moyenne (10 répétitions) Voyelles en contexte : à partir de la valeur relevée à la moitié de la durée de voyelles finales du logatome trisyllabique (10 répétitions)

Tableau 145 : Corpus sur les voyelles moyennes du français et calcul des moyennes formantiques

Les valeurs formantiques moyennes avec les écarts types se trouvent intégralement en annexe du chapitre 14 (pages 91 - 102 des annexes).

14.2.1 Voyelles moyennes en isolation et leur perception par les Français natifs

Le timbre de chaque voyelle moyenne répétée dix fois en isolation par les futures enseignantes de FLE a été vérifié via un test d'identification par dix auditeurs français natifs. Le résultat qui montre la force du lien perceptif (exprimée par la valeur du *fit index*, sur 5) entre le stimulus et la cible vocalique visée est fourni pour les locutrices T1 à T10 dans le Tableau 146 (page 312).

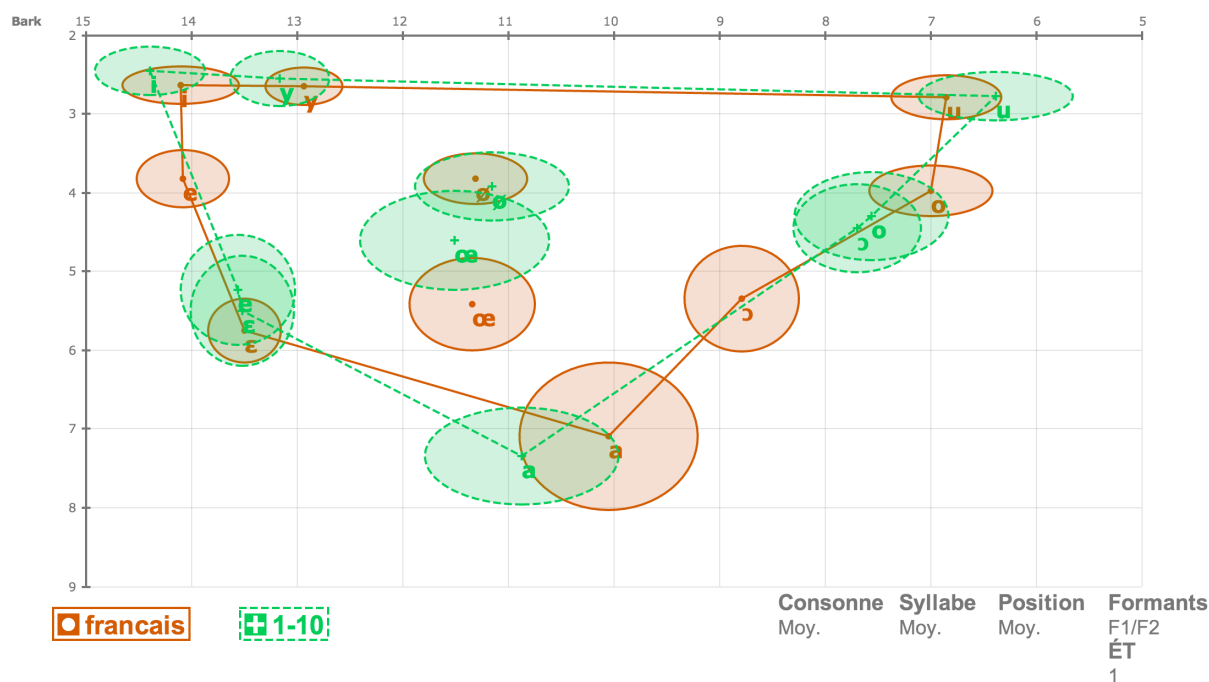
En étudiant les résultats du Tableau 146, il est possible de diviser les voyelles moyennes isolées en trois groupes, selon l'approximation possible de leur cible par des locutrices tchèques.

Alors que les voyelles [ø] et [ɛ] sont produites par la plupart des non-natives avec authenticité car identifiées comme telles par les Français natifs avec une valeur élevée du *fit index* (voir les cases grisées), les voyelles [e] et [o] ne sont généralement pas maîtrisées (ou produites de manière non-native) par les Tchèques mais leur production authentique est néanmoins possible (voir les cases grisées de la locutrice T9). Enfin, les voyelles [œ] et [ɔ] ne sont jamais maîtrisées avec authenticité ce qui est reflété par la basse valeur du *fit index* de leur identification en termes de la catégorie visée.

	e	ɛ	o	ɔ	ø	œ
T1	2,1	2,4	1,4	2,5	4,6	2,5
T2	0,4	4,2	2,9	0,9	4	1,9
T3	0,5	4,2	2,6	2,2	2,1	3,5
T4	1,3	4	3,4	0,6	4,6	3,2
T5	0	4,6	3,2	0,6	4,1	2,5
T6	0,04	4,6	0,7	2,6	4,6	0,7
T7	1	3,2	3,3	1,5	4,6	2,3
T8	0,2	3,2	2,8	1,7	4,1	1,4
T9	4,8	4,1	4,6	3	4,4	0,7
T10	0,5	4,6	2,6	2,1	3,4	2,2

Tableau 146 : Le *fit index* de l'identification des voyelles moyennes du français produites par les locutrices tchèques T1 à T10 (colonne 1), 10 répétitions par voyelle, perçues par 10 Français natifs. La valeur du *fit index* va de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte). Les voyelles produites avec authenticité (avec *fit index* supérieur à 4) sont grisées

La Figure 123 illustre la position acoustique moyenne (F1/F2, F2/F3 en Bark) des voyelles isolées produites par les dix Tchèques (trait pointillé, 10 répétitions par locutrice) par rapport à la "référence" (trait plein, 10 Françaises natives* 4 répétitions). Nous remarquons que parmi les six voyelles moyennes, seules les voyelles [ø] et [ɛ] (trait pointillé) sont en général produites à priori authentiquement car elles se trouvent à moins d'un écart type de la "référence" (trait plein). Les valeurs moyennes exactes se trouvent en annexe du chapitre 14 (page 91 des annexes).



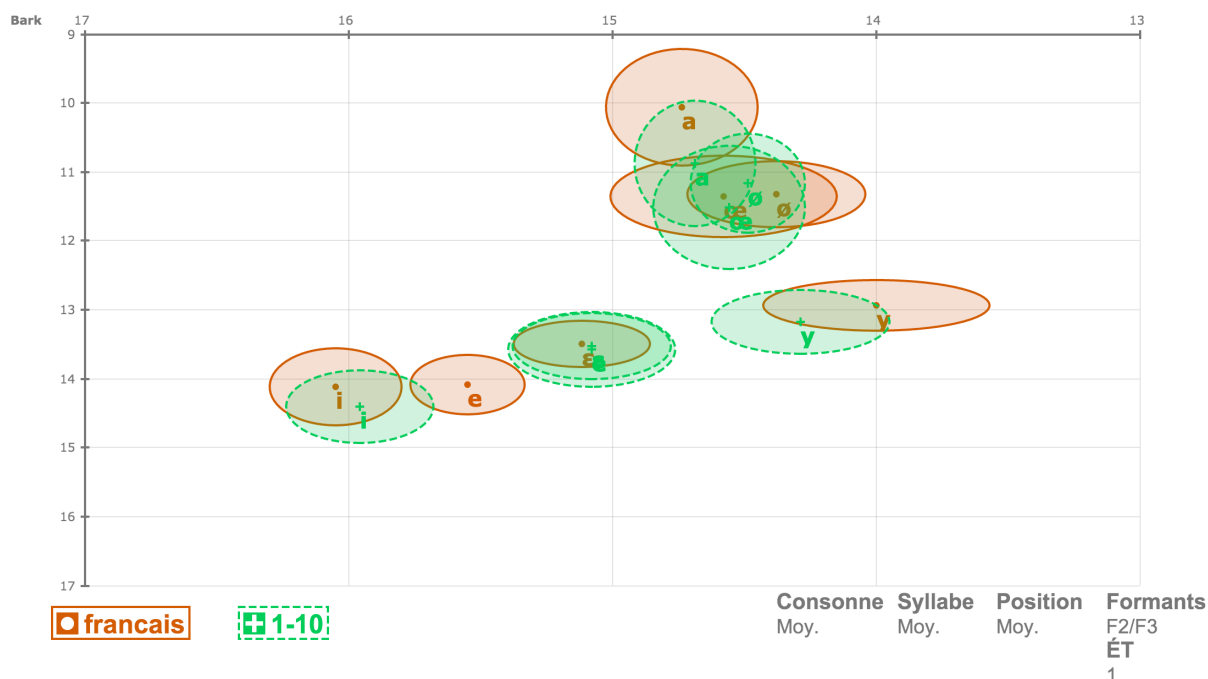


Figure 123 : Triangle F1/F2 (1^{re} image) et F2/F3 (2^e image) des voyelles isolées du français (en Bark) produites par 10 Tchèques (trait pointillé, 10 répétitions) et 10 Françaises natives (trait plein, 4 répétitions). Les voyelles [ø] et [ɛ] des Tchèques se trouvent à moins d'un écart type de la "référence"

14.2.1.1 [ø] et [ɛ] isolés, maîtrise avec authenticité courante

14.2.1.1.1 Cible [ø]

La voyelle isolée du français [ø] est *nouvelle* pour les tchécophones et il serait donc possible, selon le SLM, de l'acquérir authentiquement.

Le Tableau 147 indique le *fit index* de l'identification par les Français natifs du [ø] en termes de /ø/ ou /œ/ (puis chez la locutrice T4 et T9 en termes de /y/ également) produits par les locutrices tchèques T1 à T10 (10 répétitions par voyelle).

Stimulus	Réponse	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
ø	ø	4,6	4	2,1	4,6	4,1	4,6	4,6	4,1	4,4	4,3
	œ	0,1	0,7	2,1	/	0,5	0,1	0,1	0,4	/	0,4
	y				0,1					0,3	

Tableau 147 : Le *fit index* de l'identification du [ø] en termes de /ø/ ou /œ/ (ou /y/). Les voyelles ont été produites par les locutrices tchèques T1 à T10, 10 répétitions, identifiées par 10 Français natifs. Le *fit index* s'étend de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte). Le *fit index* égal ou supérieur à 4 (réflétant l'atteinte authentique de la cible) est en gras

Selon les résultats du Tableau 147, la voyelle [ø] est maîtrisée le plus souvent avec authenticité car neuf locutrices tchèques sur dix (toutes à l'exception de la locutrice T3) produisent une qualité vocalique qui est majoritairement catégorisée (avec un *fit index* de 4 et plus) par les natifs comme /ø/. Le spectrogramme du [ø] isolé produit par ces neuf locutrices se trouve à la Figure 124 (page 314).

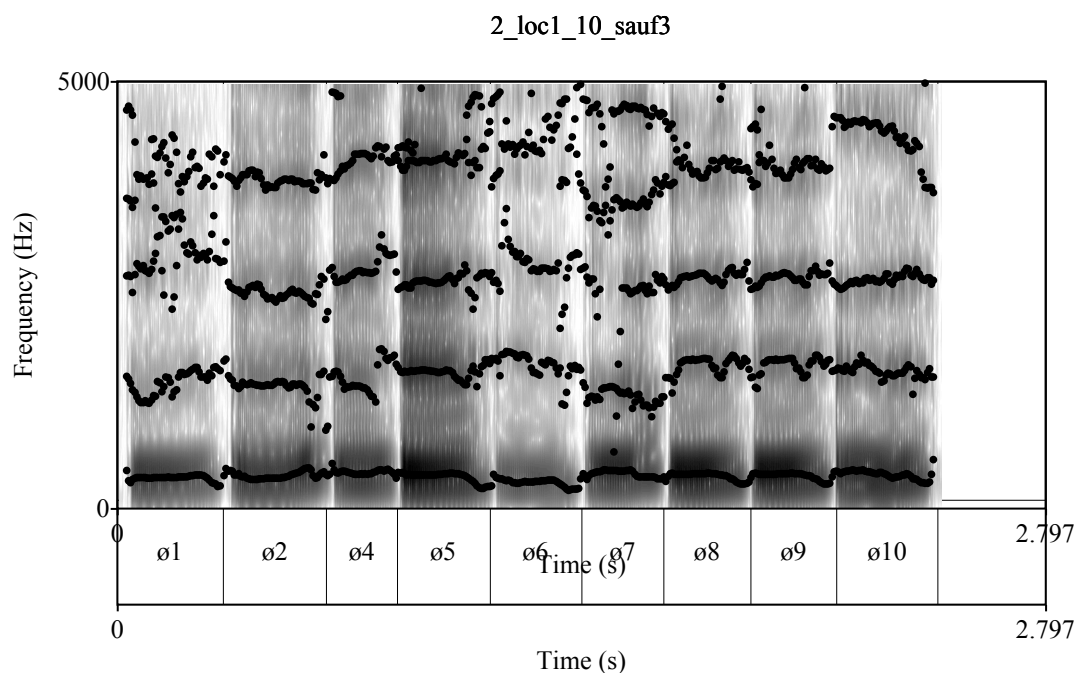


Figure 124 : Spectrogramme du [ø] isolé produit par les locutrices tchèques 1, 2, 4, 5, 6, 7, 8, 9 et 10, majoritairement identifié par les Français natifs comme tel (avec un *fit index* de 4 et plus sur 5)

Les valeurs formantiques F1, F2 et F3 (en Hertz, prises à la moitié de la durée de la voyelle) du [ø] produit par les locutrices tchèques T1, T2, T4, T5, T6, T7, T8, T9 et T10 se trouvent dans le Tableau 148.

ø	T1	T2	T4	T5	T6	T7	T8	T9	T10
F1	375	355	406	398	318	445	412	414	428
F2	1580	1443	1435	1607	1738	1355	1747	1756	1674
F3	2806	2502	2789	2642	2890	2584	2761	2758	2683

Tableau 148 : Valeurs de F1, F2, F3 (en Hz) de [ø] isolé produits par les locutrices tchèques 1, 2, 4, 5, 6, 7, 8, 9 et 10. Valeurs relevées à la moitié de la durée vocalique. La valeur minimale et maximale pour chaque formant est en gras

Le Tableau 148 révèle que le premier formant varie selon la locutrice de 318 Hz - 445 Hz, le deuxième formant de 1355 Hz – 1756 Hz et le troisième formant de 2502 Hz – 2890 Hz. Malgré la grande variation des F2 et F3, les dix productions sont catégorisées comme /ø/ avec une valeur du *fit index* de 4 ou plus sur 5. Ce résultat est en accord avec l'identification des sons synthétisés avec Klatt (en 13.1.5) qui indique une large plage de variation de F2 des sons identifiés en termes de /ø/.

Le spectrogramme de la Figure 125 illustre ensuite les dix répétitions du [ø] isolé de la locutrice T3 qui a été largement confondu par les natifs avec /œ/. En effet, dans 50 réponses sur 100 (10 répétitions*10 auditeurs), les stimuli ont été catégorisés en tant que /ø/ avec une note de 4,2 sur 5 et dans les 50 réponses restantes, ils ont été catégorisés en tant que /œ/ avec une note de 4,1 sur 5.

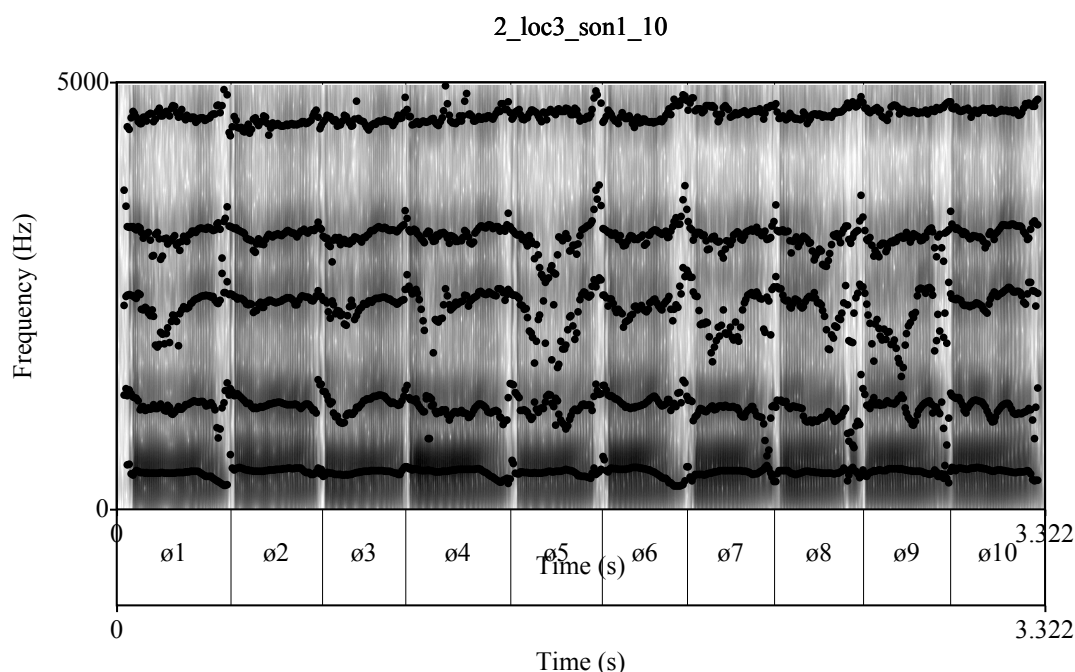


Figure 125 : Spectrogramme du [ø] isolé produit 10 fois par la locutrice tchèque T3

Les valeurs formantiques F1, F2 et F3 (en Hertz) des dix répliques du [ø] produit par la locutrice tchèque T3 se trouvent dans le Tableau 149.

T3		ø 1	ø 2	ø 3	ø 4	ø 5	ø 6	ø 7	ø 8	ø 9	ø 10
F1		447	470	454	469	431	484	458	449	480	491
F2		1258	1255	1308	1256	1264	1225	1223	1245	1251	1281
F3		2415	2425	2433	2440	2467	2477	2498	2485	2488	2424
Perçu /ø/	Nb/10	7	5	5	6	5	4	7	4	3	4
	Note	4	4	4,8	3,8	4,4	4,8	4	4	4,6	3,8
Perçu /œ/	Nb/10	3	5	5	4	5	6	3	6	7	6
	Note	4,3	4,6	3,8	4	3,6	4,3	5	3,8	4	4,2

Tableau 149 : Valeurs de F1, F2, F3 (en Hz) de 10 occurrences du [ø] isolé produit par la locutrice tchèque T3 et identifiées par 10 Français natifs. Valeurs relevées à la moitié de la durée vocalique. La valeur minimale et maximale pour chaque formant est en gras

Le Tableau 149 montre que le premier formant du [ø] produit dix fois par la locutrice tchèque T3 varie de 431 Hz à 491 Hz, le deuxième formant de 1223 Hz à 1308 Hz et le troisième formant de 2415 Hz à 2498 Hz. Les dix productions ont ainsi un patron formantique semblable et relativement stable. Alors que F1 est parfois plus élevé que le F1 maximal des neuf autres locutrices, le F2 ainsi que le F3 sont toujours plus bas que le F2 et F3 minimal des autres locutrices. Les dix qualités vocaliques produites sont alors intermédiaires entre le [ø] et le [œ] des natifs et leur identification dépend fortement des auditeurs qui - comme le met en évidence la note globale - sont souvent sûrs de leur choix (les notes s'étendent de 3,6 à 4,8 selon le stimulus). En revanche aucun stimulus n'est systématiquement catégorisé comme l'une ou l'autre voyelle. Les triangles vocaliques F1/F2 et F2/F3

(en Bark) de la Figure 126 illustrent la position moyenne du [ø] isolé produit par la locutrice T3 (en traits pointillés) par rapport à la “référence” (en trait plein).

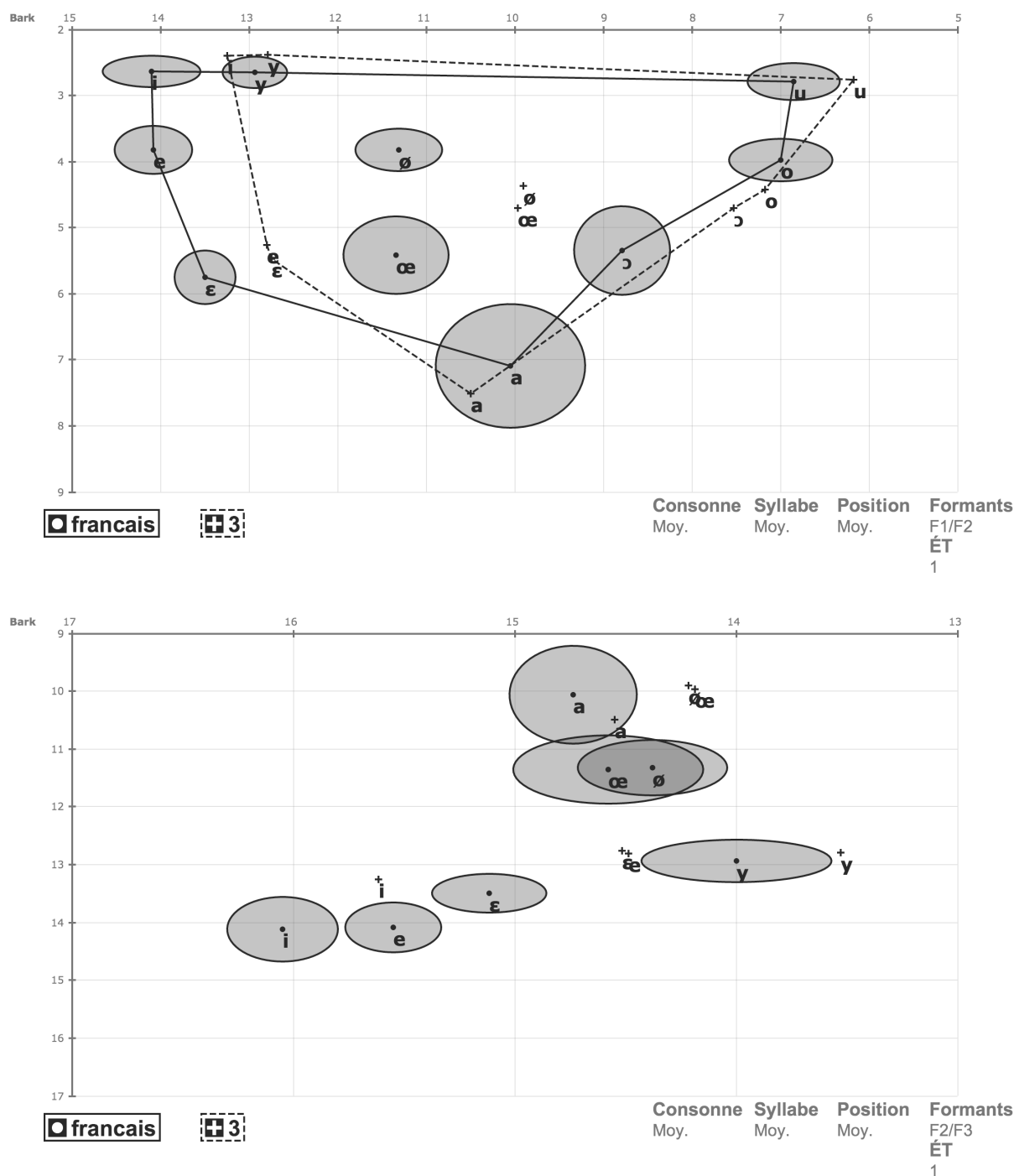


Figure 126 : Triangles F1/F2 et F2/F3 (en bas) des voyelles isolées du français (en Bark) prononcées 10 fois par la locutrice T3 (en trait pointillé) et 4 fois par 10 Françaises natives non-méridionales (en trait plein). L'écart type affiché est de un

En résumé, la cible vocalique [ø] en isolation est globalement acquise avec authenticité ce qui est en concordance avec le SLM qui prévoit la possibilité de parfaite acquisition de voyelles *nouvelles*.

14.2.1.1.2 Cible [ɛ]

La voyelle isolée du français [ɛ] est *similaire* à la voyelle tchèque [ɛ] quoique cette dernière se trouve dans l'espace acoustique F1/F2 plus bas, avec F1 en moyenne plus élevé et F2 plus bas. Ainsi, si cette voyelle est produite par simple transfert, elle ne devrait pas être confondue avec le [e] français.

Le Tableau 150 indique le *fit index* de l'identification par les Français natifs de [ɛ] en termes de /ɛ/ ou /e/ (puis chez la locutrice T5 en termes de /œ/) produits par les locutrices tchèques T1 à T10 (10 répétitions par voyelle). Nous remarquons que la confusion non attendue avec le /e/ se produit très rarement ou jamais chez les locutrices T2 à T10.

Stimulus	Réponse	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
ɛ	ɛ	2,4	4,2	4,2	4	4,6	4,6	3,2	3,2	4,1	4,6
	e	1,6	0,3	0,1	0,2		0,04	0,8	0,4	0,4	/
	œ					0,2					

Tableau 150 : Le *fit index* de l'identification du [ɛ] en termes de /ɛ/, /e/ ou /œ/. Les voyelles ont été produites par les locutrices tchèques T1 à T10, 10 répétitions, identifiées par 10 Français natifs. Le *fit index* s'étend de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte). Le *fit index* égal ou supérieur à 4 (réflétant l'atteinte authentique de la cible) est en gras

Selon le Tableau 150, la voyelle isolée [ɛ] des locutrices T2, T3, T4, T5, T6, T9 et T10 est facilement identifiable car catégorisée comme telle par les Français natifs avec un *fit index* de 4 et plus sur 5. Le spectrogramme du [ɛ] isolé produit par ces sept locutrices se trouve à la Figure 127.

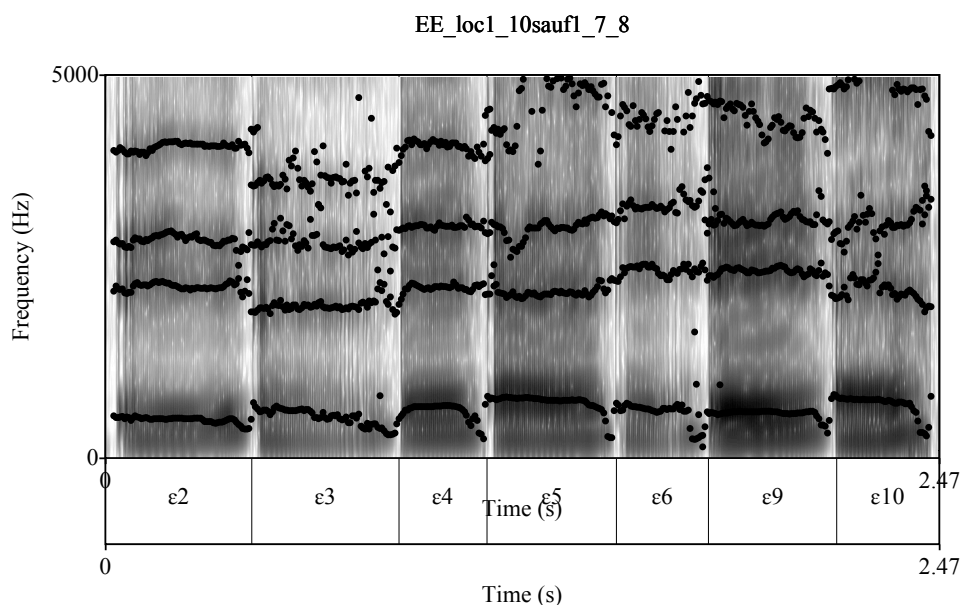


Figure 127 : Spectrogramme du [ɛ] isolé produits par les locutrices tchèques T2, T3, T4, T5, T6, T9 et T10, majoritairement identifié par les Français natifs comme tel (avec un *fit index* de 4 et plus sur 5)

Les valeurs formantiques F1, F2 et F3 (en Hertz) du [ɛ] produit par les locutrices tchèques T2, T3, T4, T5, T6, T9 et T10 se trouvent dans le Tableau 151. Les valeurs minimale et maximale pour chaque formant sont en gras. La variation des formants est grande selon la locutrice : le F1 varie de 512 à 758 Hz, le F2 de 1970 à 2533 Hz et le F3 de 2782 à 3275 Hz. Malgré la grande variation acoustique, les dix répliques du [ɛ] des sept locutrices Tchèques ont été identifiées en termes de la catégorie visée par les Français natifs avec une forte note de confiance ce qui montre que cette voyelle est produite avec authenticité.

ɛ	T2	T3	T4	T5	T6	T9	T10
F1	512	639	686	758	647	609	741
F2	2296	1970	2243	2150	2533	2422	2249
F3	2937	2782	3038	3005	3275	3038	3017

Tableau 151 : Valeurs de F1, F2, F3 (en Hz) de [ɛ] isolé produits par les locutrices tchèques 2, 3, 4, 5, 6, 9 et 10. Valeurs relevées à la moitié de la durée vocalique. La valeur minimale et maximale pour chaque formant est en gras

Les sept locutrices produisent alors un timbre qui est systématiquement identifié par les Français natifs en termes de /ɛ/ (avec un *fit index* de 4 et plus). Les trois locutrices restantes (T1, T7 et T8) produisent une qualité vocalique qui est majoritairement identifiée en termes de /ɛ/ mais une ou deux productions de chaque locutrice sont également majoritairement interprétées en termes de /e/, puis quelques productions intermédiaires renvoient aux catégories /œ/, /ø/ ou /e/. Les spectrogrammes, les valeurs formantiques et le *fit index* d'identification de chacune des dix répliques du [ɛ] de ces trois locutrices se trouvent commentés en annexe du chapitre 14, pages 113 - 116 des annexes. Les productions des trois locutrices montrent que les voyelles *similaires* ne sont pas toujours produites par simple transfert de la catégorie maternelle la plus proche.

La voyelle française [ɛ] est produite avec authenticité par sept tchécophones. La fluctuation des caractéristiques spectrales du [ɛ] chez les trois tchécophones restantes peut s'expliquer par l'apprentissage et l'application partielle de la loi de position qui régit de plus en plus la distribution des voyelles moyennes en français actuel. Dans ce sens, la fermeture vocalique des voyelles moyennes en isolation ne doit pas être considérée comme de l'échec dans l'apprentissage phonétique. Pour vérifier si les tchécophones maîtrisent la voyelle mi-ouverte [ɛ], il faut observer sa production et sa perception en syllabe finale CVC ce qui fait l'objet de la section 14.2.2.

14.2.1.2 [e] et [o] isolés, maîtrise (avec authenticité) rare mais possible

14.2.1.2.1 Cible [e]

La voyelle isolée du français [e] est très peu *similaire* au [ɛ] tchèque. Cependant du fait de son écriture par le graphème <e> qui correspond en tchèque toujours à la voyelle mi-ouverte [ɛ], un transfert négatif du [ɛ] tchèque est attendue.

Le Tableau 152 indiquant le *fit index* de l'identification par les Français natifs du [e] isolé en termes des /e/ ou /ɛ/ produits par les locutrices tchèques T1 à T10 (dix répétitions par voyelle) met en évidence que cette voyelle n'est pas systématiquement identifiée par les natifs comme telle mais que sa production authentique est néanmoins possible (reflétée par une grande valeur du *fit index* de l'identification du [e] de la locutrice T9).

Stimulus	Réponse	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
e	e	2,1	0,4	0,5	1,3	0	0,04	1	0,2	4,8	0,5
	ɛ	1,9	4	3,4	2,8	4,6	4,5	2,7	3,6	0,03	3,9

Tableau 152 : Le *fit index* de l'identification du [e] en termes de /e/ ou /ɛ/. Les voyelles ont été produites par les locutrices tchèques T1 à T10, 10 répétitions, identifiées par 10 Français natifs. Le *fit index* s'étend de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte). Le *fit index* égal ou supérieur à 4 (reflétant l'atteinte authentique de la cible) est en gras

Selon le Tableau 152, le [e] mi-fermé en isolation des locutrices tchèques T2, T3, T6, T8 et T10 est presque systématiquement interprété par les Français natifs en termes de /ɛ/ mi-ouvert (quoique pas toujours avec une grande note de qualité) et celui de la locutrice T5 est toujours identifié comme /ɛ/. Les locutrices T4 et T7 produisent ensuite une qualité vocalique qui est minoritairement interprétée comme /e/, le reste des stimuli sont identifiés de nouveau en termes de /ɛ/. La locutrice T1, elle, produit une qualité qui est majoritairement interprétée comme /e/ mais également largement confondue avec /ɛ/. Pour des raisons de fluidité de la lecture de ce travail, les spectrogrammes, les valeurs formantiques moyennes ainsi que le *fit index* de l'identification de chacune des dix répliques du [e] des locutrices T4, T7 et T1 se trouvent commentés en annexe du chapitre 14, pages 116 - 120 des annexes.

Enfin, la locutrice T9 est la seule à maîtriser complètement le [e] qu'elle produit alors avec authenticité : en effet dans 99 % des réponses (10 répétitions*10 auditeurs), le son [e] est "correctement" catégorisé, avec une note moyenne de 4,8 sur 5 (soit un *fit index* de 4,8). Il s'agit alors du meilleur lien perceptif entre un stimulus et la catégorie visée, toutes voyelles et locutrices confondues.

Le spectrogramme de la Figure 128 illustre les dix répétitions du [e] isolé qui est produit par la locutrice T9 avec authenticité. Le F1 moyen est de 440 Hz, le F2 de 2727 Hz et le F3 de 3339 Hz.

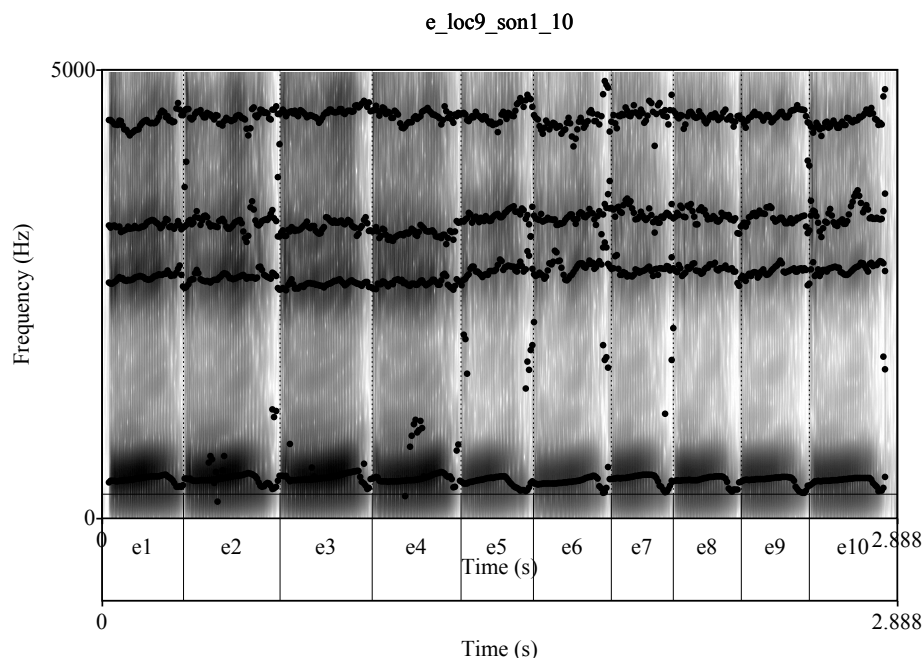
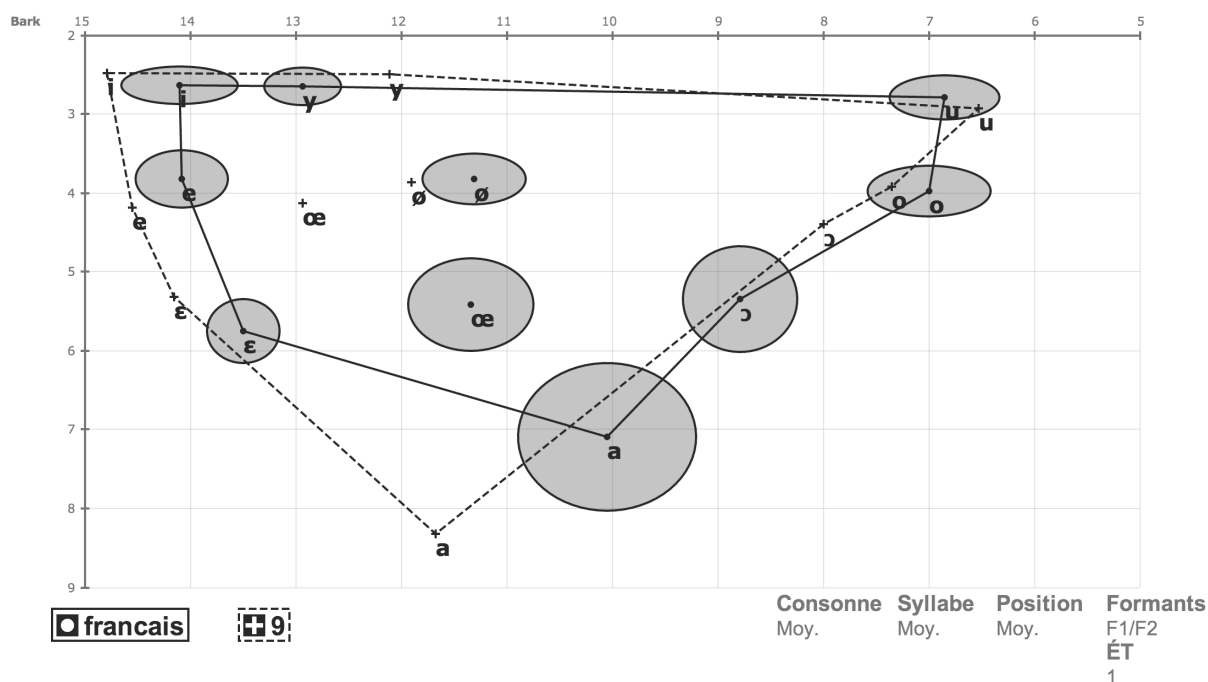


Figure 128 : Spectrogramme du [e] isolé produit 10 fois par la locutrice tchèque T9

Les triangles F1/F2 et F2/F3 (en Bark) de la Figure 129 montrent le rapprochement du [e] isolé de la locutrice T9 (10 répétitions, trait pointillé) de la "référence" (calculée à partir de 4 répliques par dix Françaises natives, trait plein). L'écart type affiché est de un.



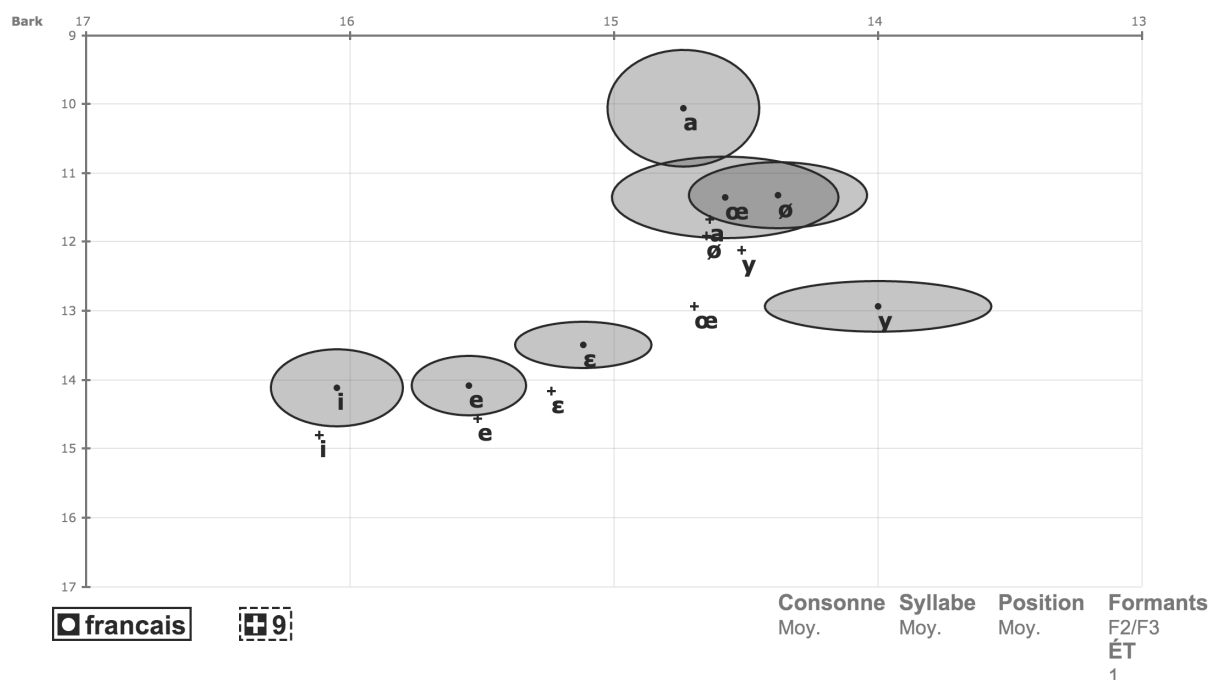


Figure 129 : Triangles F1/F2 et F2/F3 des voyelles isolées du français (en Bark) prononcées 10 fois par la locutrice T9 (en trait pointillé) et 4 fois par 10 Françaises natives non-méridionales (en trait plein). L'écart type affiché est de un

Ainsi, la qualité vocalique [e] n'est généralement pas maîtrisée. Six locutrices réalisent cette voyelle de manière à ce qu'elle soit toujours majoritairement catégorisée en termes de /e/. Trois autres locutrices ne maîtrisent pas complètement cette qualité mais montrent néanmoins un progrès : les dix productions du [e] de la locutrice T7 sont alors soit majoritairement interprétés en termes de /e/, soit interprétées en termes des deux timbres /e/ ou /ε/ même si l'identification par /e/ est accompagnée d'une meilleure note. Celles de la locutrice T1 sont ensuite soit majoritairement catégorisées comme /e/, soit comme /ε/, soit renvoient aux deux timbres sur le même pied d'égalité. L'identification en termes de /e/ est accompagnée d'une meilleure note. Il est à noter que l'existence de la catégorie mentale pour [e] a déjà été observée chez les locutrices T1 et T7 lors de la production de certaines répliques du [ε] qui tendaient vers la fermeture.

La locutrice T4 réalise ensuite des qualités vocaliques qui renvoient majoritairement soit au /e/, soit au /ε/, mais contrairement aux locutrices T7 et T1, une même production ne renvoie jamais dans la même mesure aux deux timbres. Enfin, la locutrice T9 est la seule à maîtriser systématiquement cette voyelle et elle la produit fidèlement.

La voyelle isolée [e] (souvent notée par le graphème <é>) n'est généralement pas maîtrisée et nous expliquons ce résultat par l'influence de la graphie. En effet, le graphème <é> se prononce en tchèque systématiquement avec [ε], qui est *similaire* au [ε] français. Une tchécophone réalise cependant le [e] mi-fermé avec authenticité. Ce résultat montre qu'il est possible de surmonter l'influence de la graphie qui intervient au cours de l'apprentissage phonétique formel. Ainsi, nous confirmons la validité du SLM seulement en partie car ce modèle ne prévoit pas de possibilité d'apprentissage phonétique authentique des sons *similaires* (et donc du [e]).

14.2.1.2.2 Cible [o]

La voyelle isolée du français [o] est très peu *similaire* aux [o/o:] tchèques mais du fait du symbole phonétique utilisé et du graphème <o> qui correspond en tchèque toujours au phonème [o], un transfert de [o] tchèque dans les productions en français est attendu. Comme le [o] isolé tchèque se trouve dans l'espace acoustique F1/F2 entre le [o] et le [ɔ] français, nous nous attendons à ce que les réalisations du [o] par les Tchèques soient interprétés par les Français natifs en termes des deux

catégories. La situation est donc toute différente de la voyelle tchèque [ɛ] qui par simple transfert ne peut correspondre qu'au seul timbre français [ɛ].

Le Tableau 153 indique le *fit index* de l'identification par les Français natifs du [o] isolé des locutrices tchèques T1 à T10 (10 répétitions par voyelle) en termes de /o/ ou /ɔ/.

Stimulus	Réponse	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
o	o	1,4	2,9	2,6	3,4	3,2	0,6	3,3	2,8	4,6	2,6
	ɔ	2,7	1,2	1,5	0,9	1	3,4	0,9	1,5	0,1	1,5

Tableau 153 : Le *fit index* de l'identification du [o] en termes de [o] ou [ɔ]. Les voyelles ont été produites par les locutrices tchèques T1 à T10, 10 répétitions, identifiées par 10 Français natifs. Le *fit index* s'étend de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte). Le *fit index* égal ou supérieur à 4 (reflétant l'atteinte authentique de la cible) est en gras

Le Tableau 153 montre que le timbre mi-fermé [o] n'est généralement pas maîtrisé par les Tchèques : il est en effet soit majoritairement interprété par les natifs en termes de /o/ mais confondu également avec /ɔ/ (dans le cas des productions des locutrices T2, T3, T4, T5, T7, T8, T10), soit majoritairement interprété en termes de /ɔ/ mais également confondu avec /o/ (productions des locutrices T1 et T6). Ce résultat est attendu si l'on considère que les Tchèques produisent cette voyelle par transfert de la LM. En revanche, les résultats montrent que la locutrice T9 maîtrise le [o] mi-fermé avec authenticité. Le spectrogramme des dix occurrences du [o] produit par la locutrice T9 se trouve alors à la Figure 130. Le F1 est en moyenne de 410 et le F2 de 837 Hz.

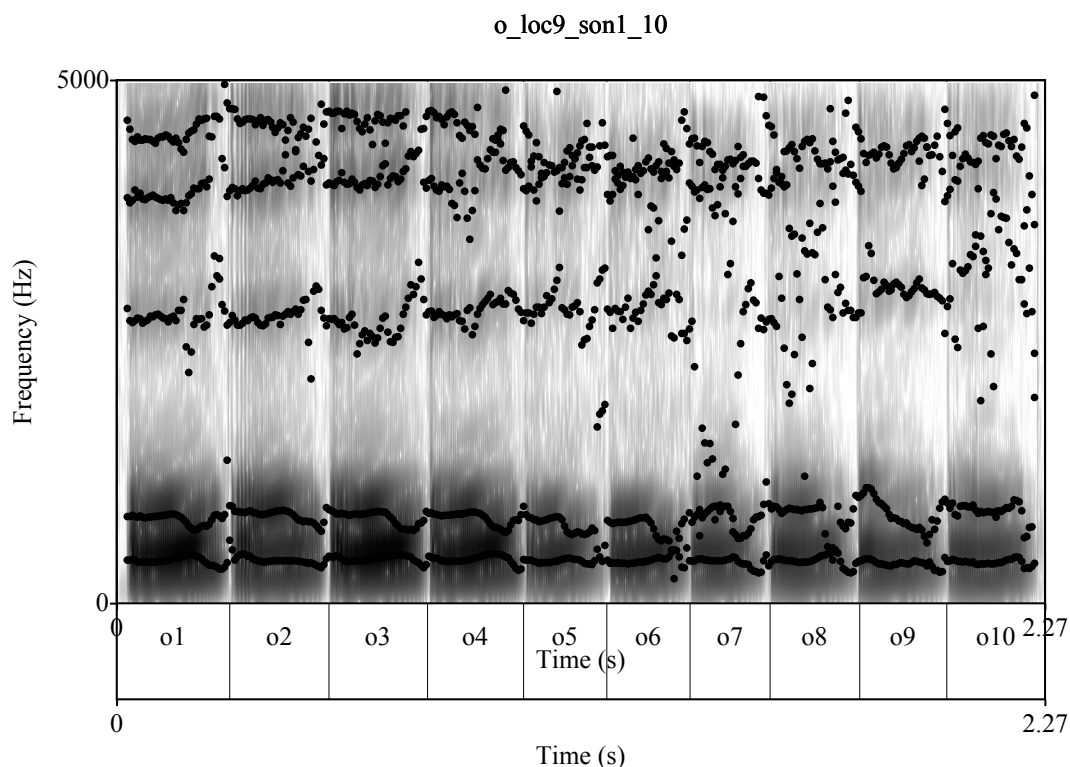


Figure 130 : Spectrogramme de dix répétitions du [o] isolé produit par la locutrice tchèque T9

Les dix occurrences du [o] de la locutrice T9 sont ensuite illustrées dans le triangle vocalique F1/F2 (en Bark) de la Figure 131 (page 322). Nous y remarquons la stabilité du timbre vocalique.

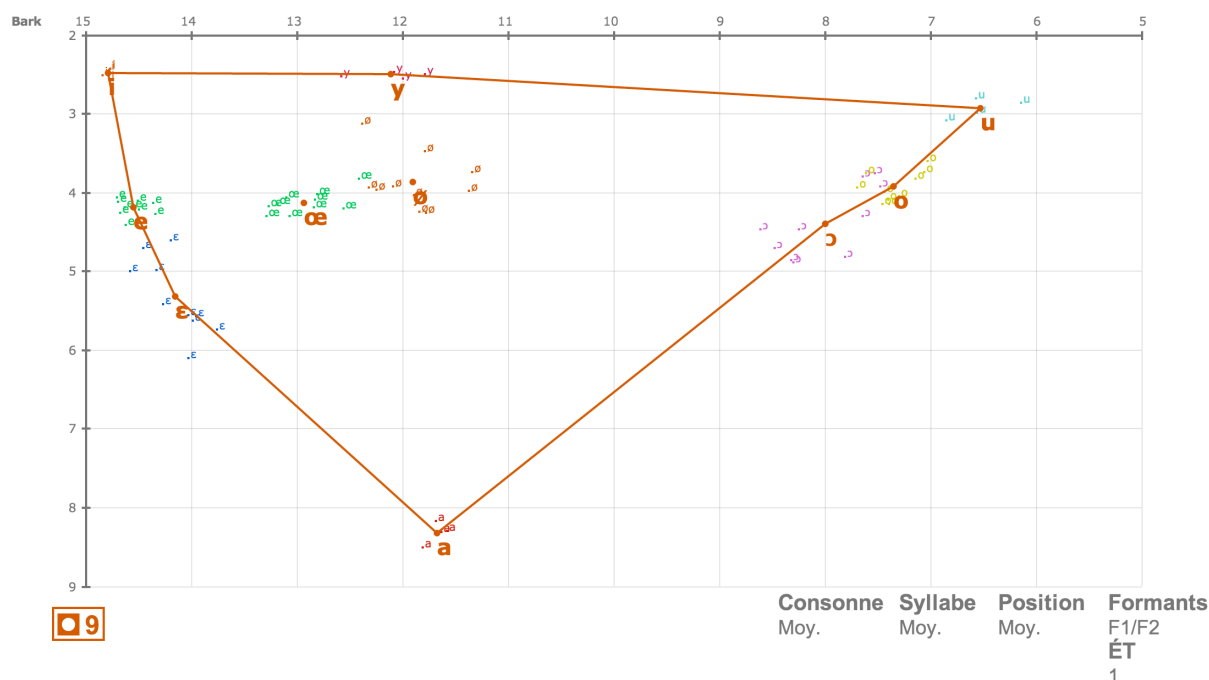


Figure 131 : Triangle F1/F2 (en Bark) des voyelles isolées du français prononcées 10 fois par la locutrice T9. Remarquons la stabilité dans la production du timbre vocalique [o]

Les dix répliques du [o] des locutrices T3, T4, T7, T1 et T6 ont également été analysées et les spectrogrammes, les valeurs formantiques des locutrices ainsi que les *fit index* de l'identification de chacun des timbres se trouvent en annexe du chapitre 14, pages 120 - 128 des annexes. Le résultat de l'analyse montre que pour la perception du timbre visé /o/, les deux formants F1 et F2 doivent être bas. En effet lorsque F2 est élevé, une valeur même très basse de F1 ne saurait pas provoquer la perception de /o/. Ce résultat est en accord avec l'identification des sons vocaliques générés avec Klatt (en 13.5.1).

Pour résumer, la voyelle [o] des futures enseignantes tchèques est souvent réalisée par transfert de la LM : elle se trouve dans l'espace vocalique à mi-chemin entre le [o] et le [ɔ] des natives. Sa position acoustique explique alors pourquoi elle est confondue par les Français natifs avec /ɔ/. Toutefois, une locutrice tchèque réalise cette voyelle *similaire* de façon authentique, résultat qui n'est pas envisagé par le SLM.

14.2.1.3 [œ] isolé, maîtrise (sans authenticité) rare mais possible

Bien que la voyelle isolée du français [œ] soit « symboliquement » et acoustiquement *nouvelle* pour les tchécophones, elle est perceptivement *similaire* aux [ɛ/ɛ:] tchèques. Sur la base du SLM, l'acquisition authentique de cette voyelle est donc exclue.

En effet, la voyelle isolée [œ] des locutrices tchèques est identifiée par les Français natifs en termes de la catégorie visée avec une faible valeur du *fit index* (elle varie de 0,7 à 3,5 selon la locutrice) et elle est fréquemment confondue avec la qualité mi-fermée /ø/ (qui est une voyelle *nouvelle*).

Le Tableau 154 indique le *fit index* de l'identification par les Français natifs du [œ] des locutrices tchèques T1 à T10 en termes des /œ/ ou /ø/. Le [œ] est soit majoritairement catégorisé comme /œ/ mais également confondu avec /ø/ (productions des locutrices T1, T3, T4, T5, T7 et T10), soit majoritairement catégorisé comme /ø/ mais également comme /œ/ (productions des locutrices T2, T6, T8 et T9).

Le Tableau 154 montre que même si aucune locutrice ne maîtrise la cible vocalique [œ] avec authenticité, elle est maximalelement rapprochée par la locutrice T3 dont le [œ] isolé est “correctement” identifié dans 88 % des réponses avec la note de 3,9 sur 5 (soit un *fit index* de 3,5), puis par la locutrice T4 dont le [œ] isolé est “correctement” identifié dans 79 % des réponses avec la note de 4 sur 5 (soit un *fit index* de 3,2).

Stimulus	Réponse	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
œ	œ	2,5	1,9	3,5	3,2	2,5	0,7	2,3	1,4	0,7	2,2
	ø	1,9	2,5	0,6	0,6	1,6	3,8	1,9	1,9	3	1,2

Tableau 154 : Le *fit index* de l'identification du [œ] en termes de /œ/ ou /ø/. Les voyelles ont été produites par les locutrices tchèques T1 à T10, 10 répétitions, identifiées par 10 Français natifs. Le *fit index* s'étend de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte). Notons que le [œ] des seules locutrices T3 et T4 a été majoritairement catégorisé comme tel et peu confondu avec /ø/

Les dix productions du [œ] de la locutrice T3 qui se rapproche le plus systématiquement de la “référence” sont illustrées dans le spectrogramme de la Figure 132.

9_loc3_son1_10

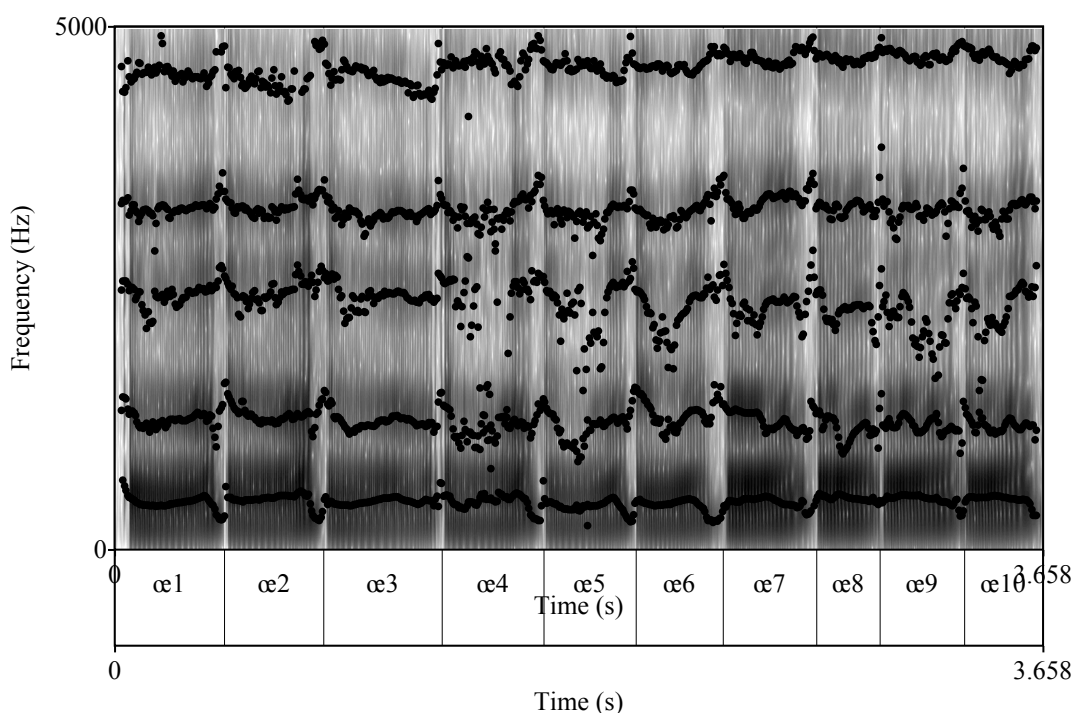


Figure 132 : Spectrogramme de dix répétitions de [œ] isolé produites par la locutrice tchèque T3

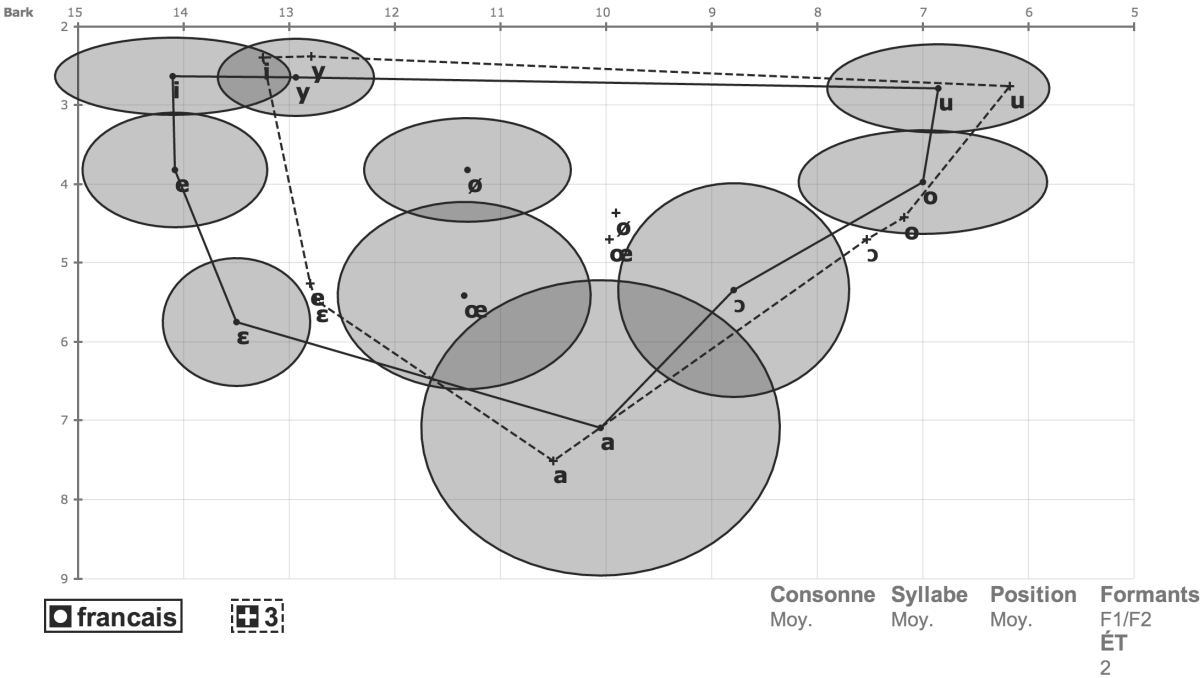
Les valeurs formantiques F1, F2 et F3 (en Hertz) des dix occurrences du [œ] produit par la locutrice tchèque T3 et les catégories auxquelles les productions renvoient se trouvent dans le Tableau 155 (page 324). Les valeurs minimale et maximale pour chaque formant est en gras.

Le Tableau 155 révèle que le F1 des dix répliques du [œ] produites par la locutrice T3 varie de 446 Hz à 541 Hz, le F2 de 1194 Hz à 1331 Hz et le F3 de 2314 Hz à 2538 Hz. Les dix stimuli sont majoritairement identifiés par les Français natifs en termes de la catégorie visée /œ/ et la note accordée varie de 3,6 à 4,3 suivant le stimulus. De cette manière, la locutrice T3 maîtrise la voyelle isolée [œ] mais elle ne la produit pas avec authenticité. Il est intéressant à noter que le son 6 produit avec F2 maximal est identifié par quatre auditeurs en termes du /ø/ avec une note forte de 4,8 sur 5.

T3		æ 1	æ 2	æ 3	æ 4	æ 5	æ 6	æ 7	æ 8	æ 9	æ 10
F1		446	516	483	541	485	480	483	475	521	491
F2		1247	1261	1228	1204	1232	1331	1307	1194	1264	1259
F3		2432	2460	2435	2538	2498	2420	2390	2314	2368	2426
Perçu /æ/	Nb/10	10	8	9	10	9	6	9	8	10	9
	Note	3,6	4,1	3,6	3,9	3,9	3,5	3,8	4	4,2	4,3
Perçu /ø/	Nb/10		2	1		1	4	1	2		1
	Note		4,5	5		5	4,8	5	4		5

Tableau 155 : Valeurs de F1, F2, F3 (en Hz) et identification par 10 auditeurs français natifs de [æ] isolé produit par la locutrice tchèque T3 (10 répétitions). Valeurs relevées à la moitié de la durée vocalique. La valeur minimale et maximale pour chaque formant est en gras

Les triangles vocaliques F1/F2 et F2/F3 (en Bark) de la Figure 133 illustrent la position moyenne du [æ] produit par la locutrice T3 (en trait pointillé) par rapport à la “référence” (en trait plein) où les ellipses de dispersion sont tracées à deux écarts types. Il est à noter que malgré la position acoustique du [æ] produit par la locutrice T3 à plus de deux écarts types de la “référence”, ce dernier est systématiquement identifié par les Français natifs en termes de la catégorie visée. En revanche sa production n’est pas authentique comme reflétée par la note de qualité accordée par les auditeurs français.



14. Voyelles moyennes du français chez des futures enseignantes de FLE de LM tchèque

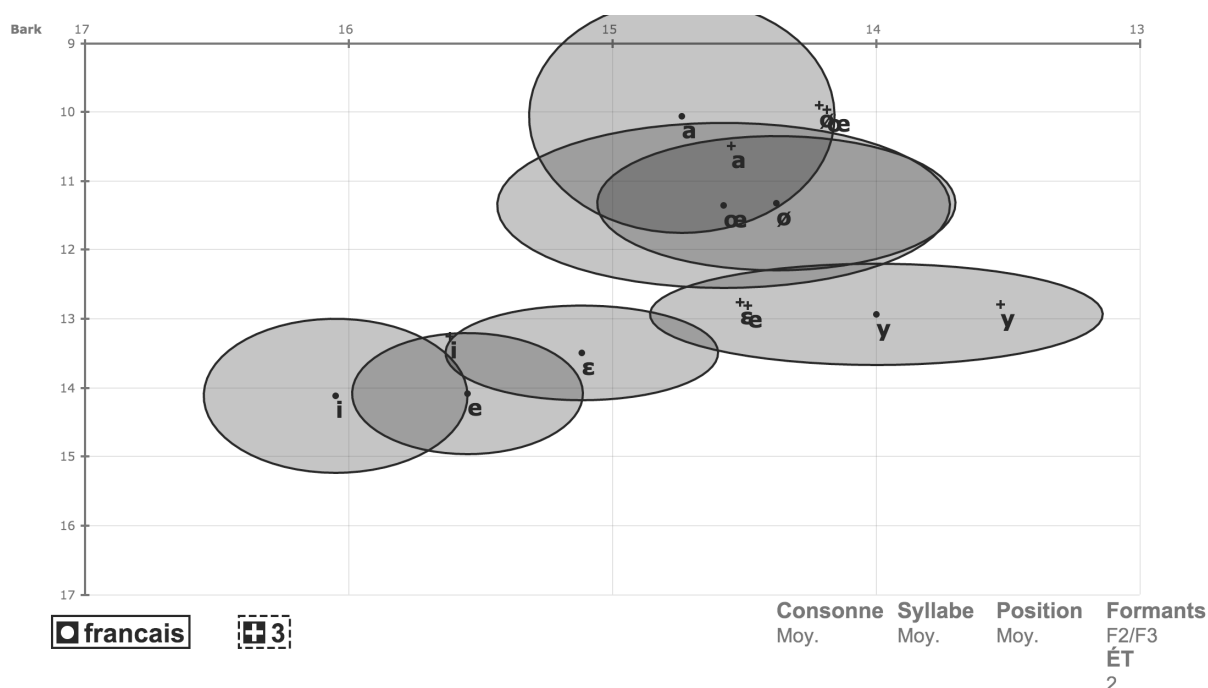


Figure 133 : Triangles F1/F2 et F2/F3 (en bas) des voyelles isolées du français (en Bark) prononcées 10 fois par la locutrice T3 (en trait pointillé) et 4 fois par 10 Françaises natives non-méridionales (en trait plein). L'écart type affiché est de deux

La locutrice T4 produit également des qualités vocaliques qui sont majoritairement identifiés en termes de /œ/. Le spectrogramme des dix occurrences du [œ] se trouve à la Figure 134.

9_loc4_son1_10

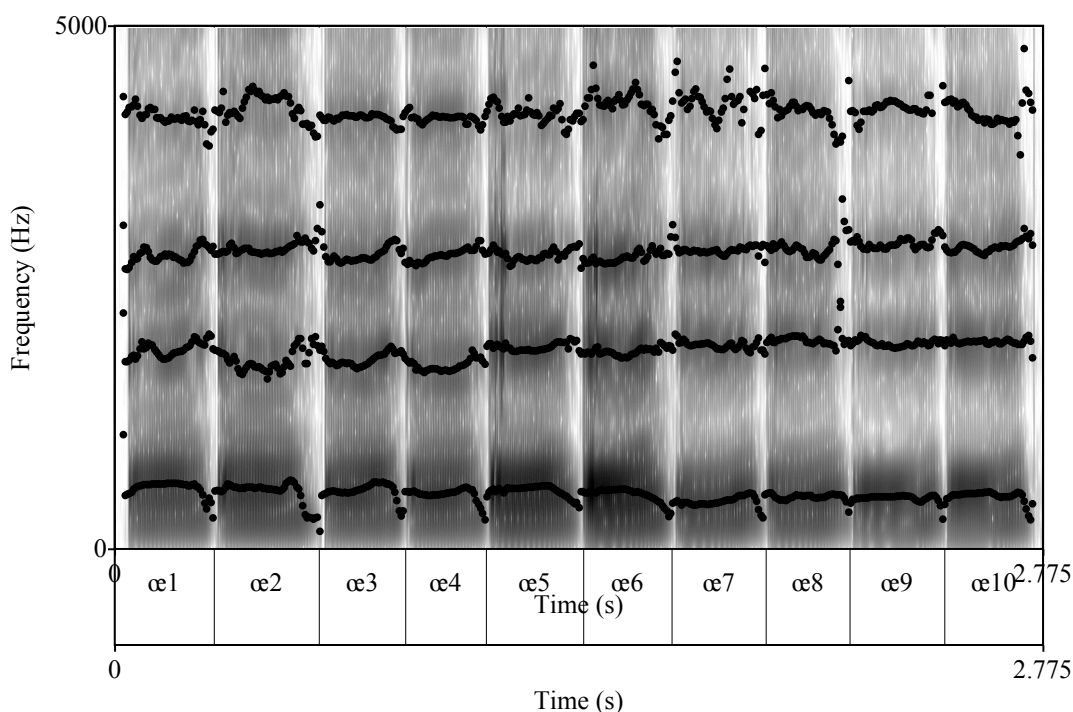


Figure 134 : Spectrogramme de dix répétitions de [œ] isolé produit par la locutrice tchèque T4

Les valeurs formantiques F1, F2 et F3 (en Hertz) des dix occurrences du [œ] produites par la locutrice T4 et les catégories auxquelles les productions renvoient se trouvent dans le Tableau 155 de la page 324. Les valeurs minimale et maximale pour chaque formant sont en gras.

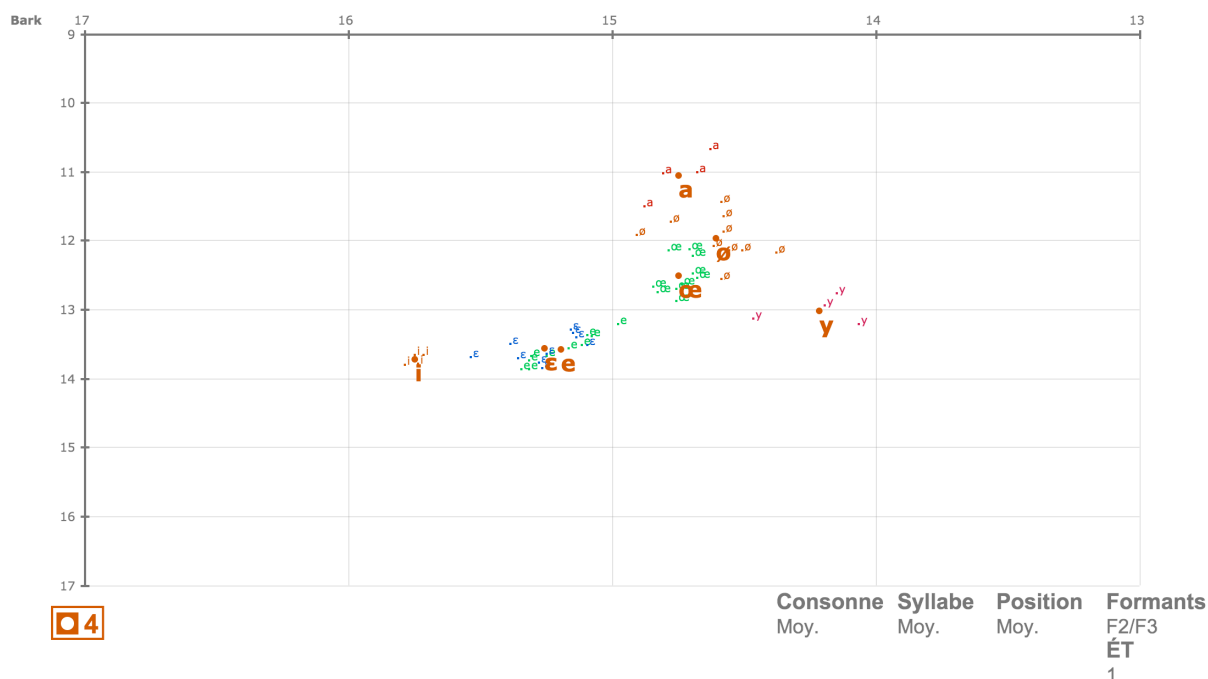


Figure 135 : Triangles F1/F2 et F2/F3 (en bas) des voyelles isolées du français (en Bark) prononcées 10 fois par la locutrice 4

Les dix répliques du [œ] réalisées par les locutrices T6 et T1 sont ensuite analysées en annexe du chapitre 14, pages 128 - 130 des annexes. On y trouve les spectrogrammes, les valeurs formantiques et le *fit index* d'identification de chacune des dix répliques en termes de la catégorie visée. Les résultats montrent un caractère fluctuant des sons représentant la qualité [œ] qui sont plus souvent interprétés en termes de /ø/. **Cependant l'instabilité de la qualité produite montre que le processus d'apprentissage est en cours.**

Ainsi, même si les locutrices tchèques produisent parfois une qualité vocalique qui est à l'unanimité identifiée en termes de la catégorie visée /œ/ avec une forte note, cette voyelle n'est systématiquement maîtrisée que par une locutrice (T3) qui la produit néanmoins de manière non-native (comme en témoigne la faible note de qualité attribuée par les Français natifs). Afin de catégoriser une voyelle en termes du timbre mi-ouvert /œ/, la valeur de son premier formant semble capitale (elle est plus élevée que celle de [ø]) alors que la valeur du deuxième formant détermine le niveau d'authenticité. Ce résultat est en accord avec les observations en 13.1.5 concernant l'identification par un Français natif de voyelles synthétisées avec Klatt.

Enfin, ce résultat appuie le SLM qui ne prévoit pas la possibilité d'apprentissage authentique des sons *similaires* aux sons maternels.

14.2.1.4 [ɔ] isolé, non maîtrisé

La voyelle isolée [ɔ] du français est *similaire* au [o/o:] tchèque. Cette dernière étant véritablement moyenne se réalise dans l'espace acoustique F1/F2 entre la qualité mi-fermée [o] et mi-ouverte [ɔ], comme l'on peut le voir à la Figure 38 à la page 132. Par simple transfert, le [ɔ] des Tchèques devrait donc être intermédiaire entre les deux qualités françaises et selon les auditeurs natifs devrait renvoyer à l'un ou l'autre timbre. Le SLM ne prévoit pas d'acquisition authentique pour ce son.

Comme indiqué dans le Tableau 157 de la page 328, la voyelle [ɔ] n'est jamais systématiquement identifiée comme telle : son *fit index* d'identification par les Français natifs est alors égal ou inférieur à 3 et ce résultat est attendu par SLM. Le [ɔ] est majoritairement catégorisé comme

tel chez les locutrices T1, T3, T6 et T9, puis comme /o/ chez les locutrices T2, T4, T5, T7, T8 ; la locutrice T10 produit une qualité vocalique qui est reliée par les natifs au /ɔ/ et /o/ avec le même *fit index*.

Stimulus	Réponse	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
ɔ	ɔ	2,5	0,9	2,2	0,6	0,6	2,6	1,5	1,7	3	2,1
	o	1,4	3,2	1,2	3,5	3,8	1,3	2,7	2,5	1,6	2,1

Tableau 157 : Le *fit index* de l'identification du [ɔ] en termes de /ɔ/ ou /o/. Les voyelles ont été produites par les locutrices tchèques T1 à T10, 10 répétitions, perçues par 10 Français natifs. Le *fit index* s'étend de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte)

Comme le révèle la valeur du *fit index* du Tableau 157, la cible [ɔ] est maximale ment rapprochée par la locutrice T9 : le timbre qu'elle produit est alors "correctement" identifié dans 65 % des réponses avec une note moyenne de 4,6 sur 5, soit un *fit index* de 3. Ses dix répliques du [ɔ] sont illustrées dans le spectrogramme de la Figure 136.

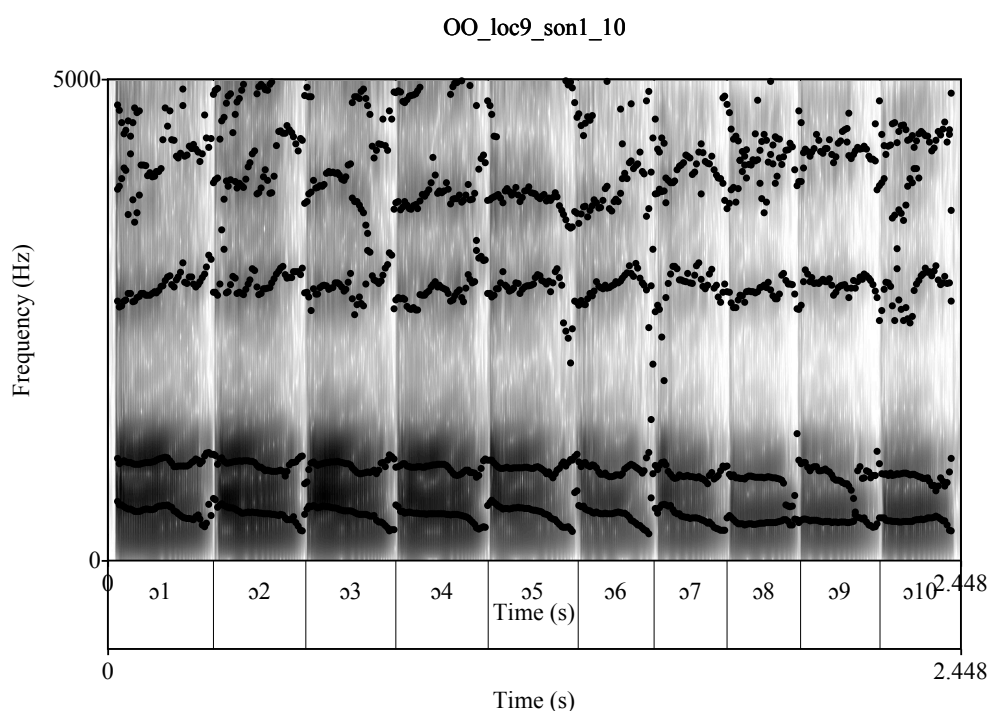


Figure 136 : Spectrogramme de dix répétitions de [ɔ] isolé produit par la locutrice tchèque T9

Les valeurs formantiques F1 et F2 (en Hertz) des dix occurrences du [ɔ] produit par la locutrice T9 et les catégories auxquelles les productions renvoient se trouvent dans le Tableau 158. La valeur minimale et maximale pour chaque formant est en gras. Nous y remarquons que le F1 du [ɔ] produit par la locutrice T9 varie de 399 Hz à 532 Hz et le F2 de 775 Hz à 1023 Hz. Les stimuli 1 à 7 sont majoritairement ou exclusivement identifiés par les auditeurs français natifs en termes de la catégorie visée /ɔ/. Ces sons ont été produits avec F1 de 442 Hz ou plus et F2 de 777 Hz ou plus. A l'inverse, les stimuli 8 à 10 sont majoritairement ou exclusivement interprétés comme /o/ : les sons se réalisent alors avec F1 de 409 Hz ou moins et F2 entre 775 Hz et 907 Hz. Il est à noter que même si les sons 7 ou 8 sont réalisés avec F2 semblable (respectivement de 777 Hz et 775 Hz), le premier (avec F1 de 442 Hz) est systématiquement identifié comme /ɔ/ et le deuxième (avec F1 de 399 Hz) comme /o/. Lorsque F2 est bas, la valeur de F1 est alors décisive dans l'identification du timbre o/ɔ. Globalement, la locutrice T9 produit des qualités vocaliques qui sont majoritairement identifiées soit comme /ɔ/ soit comme /o/ mais aucun son ne renvoie aux deux timbres sur le même pied d'égalité. De cette façon, la locutrice T9 montre un progrès dans l'approximation de la cible [ɔ] mais elle ne la maîtrise pas complètement.

T9		ɔ 1	ɔ 2	ɔ 3	ɔ 4	ɔ 5	ɔ 6	ɔ 7	ɔ 8	ɔ 9	ɔ 10
F1		532	497	522	486	529	504	442	399	403	409
F2		1023	987	1014	1005	961	891	877	875	907	895
Perçu /ɔ/	Nb/10	10	9	9	8	10	7	10	0	1	1
	Note	4,8	4,9	4,7	4,9	4,2	4,6	4,4		5	5
Perçu /o/	Nb/10		1		2		3		10	9	9
	Note		4		4		4		4,6	4,3	4,9

Tableau 158 : Valeurs de F1 et F2 (en Hz) et identification par 10 auditeurs français natifs du [ɔ] isolé produit par la locutrice tchèque T9 (10 répétitions). Valeurs relevées à la moitié de la durée vocalique. La valeur minimale et maximale pour chaque formant est en gras

Le triangle vocalique F1/F2 (en Bark) de la Figure 137 illustre les dix productions de chacune des voyelles orales du français réalisées par la locutrice T9. Nous remarquons la dispersion des dix occurrences du [ɔ] dont trois se trouvent dans la zone acoustique du [o].

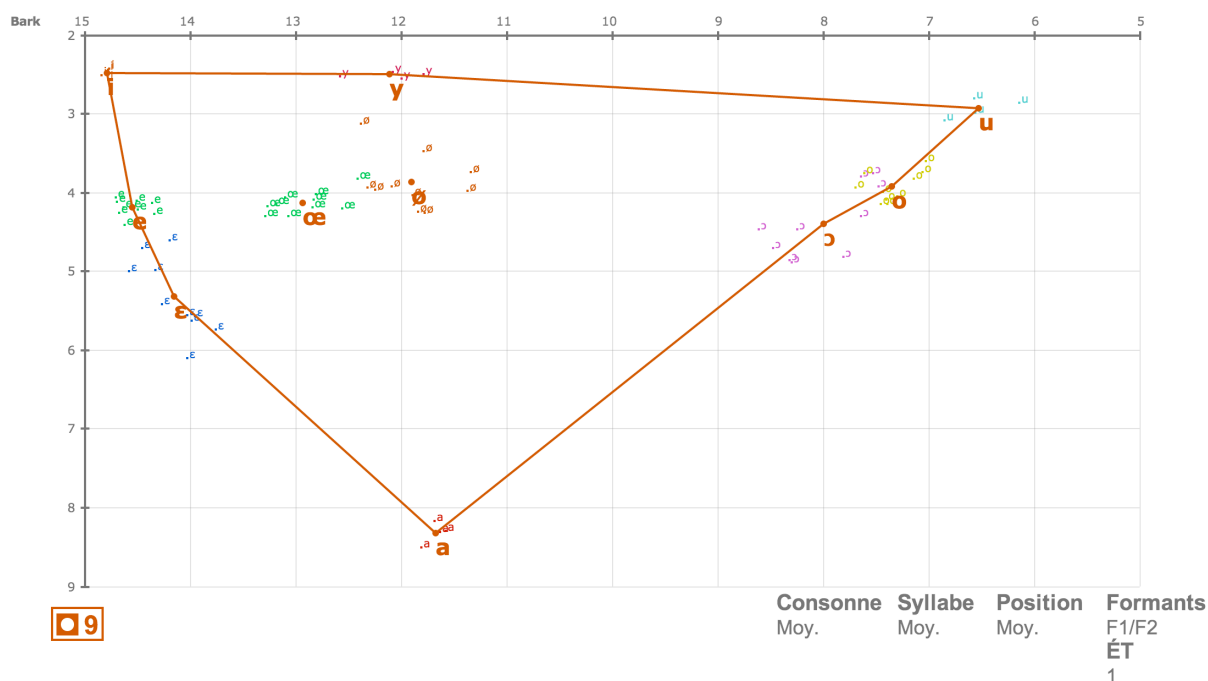


Figure 137 : Triangle F1/F2 (en Bark) des voyelles isolées du français prononcées 10 fois par la locutrice T9

Les dix répliques du [ɔ] de deux autres locutrices (T10 et T5) ont été analysées et leurs spectrogrammes, valeurs formantiques et *fit index* d'identification se trouvent commentés en annexe du chapitre 14, à la page 131 et 132 des annexes. Alors que les productions de la locutrice T10 se caractérisent par une forte fluctuation avec six répliques catégorisées majoritairement comme /ɔ/, indiquant ainsi l'approximation de la cible, les dix répliques de la locutrice T5 sont identifiées en tant que /o/.

Pour résumer, la voyelle isolée [ɔ] n'est complètement maîtrisée par aucune des futures enseignantes tchèques de FLE ce qui se reflète dans la petite valeur du *fit index* d'identification par les auditeurs français natifs. Ce résultat est en accord avec le SLM qui ne prévoit pas la possibilité d'acquisition authentique de sons *similaires*.

14.2.1.5 Productions de voyelles moyennes en isolation: Conclusion partielle

La réalisation de la qualité exacte des voyelles moyennes isolées du français par les tchécophones qui ont un système vocalique maternel à trois degrés d'aperture s'avère pour la plupart difficile et le niveau de difficulté varie selon la voyelle.

La cible la plus souvent atteinte est celle de la voyelle *nouvelle* [ø] car neuf futures enseignantes sur dix la produisent systématiquement avec authenticité, comme en témoigne la valeur élevée (égale ou supérieure à 4 sur 5) du *fit index* de son identification par les Français natifs. La cible de la voyelle *similaire* [ɛ] du fait de sa position par rapport au [ɛ] tchèque est également possible à atteindre (par simple transfert) et sept futures enseignantes la produisent avec authenticité. Si l'on considère la voyelle du français [ɛ] comme quasi *identique* au [ɛ] tchèque, les résultats appuient le SLM.

Ensuite, la réalisation du timbre des voyelles *similaires* [e] et [o] s'avère plus difficile pour les Tchèques. Probablement sous l'influence de la graphie, neuf futures enseignantes sur dix ne les maîtrisent pas (quoiqu'elles montrent un progrès dans leur apprentissage). En revanche, une locutrice (T9) les maîtrise avec fidélité ce qui n'est pas en accord avec les prédictions de SLM.

La voyelle [œ], très peu *similaire* aux [ɛ/ɛ:] tchèques n'est maîtrisée que par une future enseignante qui la produit néanmoins de manière non-native et la voyelle *similaire* [ɔ] n'est maîtrisée par aucune future enseignante. Ces résultats peuvent s'expliquer de deux façons. Soit les dix Tchèques réalisent ces deux qualités par simple transfert de la qualité maternelle la plus proche (mais on pourrait se demander alors pourquoi la locutrice T9 a su échapper au *mécanisme de classification par équivalence* pour les qualités fermées [e] et [o] et pas pour les qualités ouvertes [ɔ] et [œ]), soit en suivant la loi de position les locutrices tendent à fermer la qualité des voyelles moyennes en isolation. Si la deuxième explication est la bonne, nous ne pouvons pas considérer le résultat à propos des cibles [ɔ] et [œ] comme de l'échec mais plutôt comme de la réussite. Ainsi, pour vérifier si les Tchèques maîtrisent les qualités [ɔ] et [œ] du français en production, il est nécessaire d'étudier la réalisation des contrastes ø/œ et o/ɔ en syllabe finale fermée où les quatre phonèmes peuvent être en opposition. Ce point fera l'objet de la section suivante.

14.2.2 Contrastes entre e/ɛ, ø/œ (F1 - F3) et o/ɔ (F1, F2) en contextes nul et p, t, k, R

Dans la section présente, nous nous intéressons à la réalisation par les futures enseignantes des contrastes entre les voyelles moyennes du français prononcées en isolation et en syllabe finale CVC où C correspond aux consonnes [p, t, k, ʀ]. **En absence de tests de perception vérifiant les qualités produites dans les différents environnements consonantiques, l'objectif n'est pas de montrer si le contraste produit est perceptivement saillant mais plutôt de déterminer si les futures enseignantes réalisent des sons acoustiquement distincts représentant deux catégories différentes, ce qui témoignerait d'un apprentissage en cours.** Nous considérons que les contrastes sont réalisés (mais pas nécessairement perçus par les natifs) quand les valeurs formantiques des voyelles de chaque paire se trouvent à plus d'un écart type les unes des autres. Ensuite, plus la distance acoustique entre les deux qualités produites est grande et plus les sons représentant la même catégorie ont un caractère stable, plus il est probable que la locutrice détient deux catégories mentales différentes. Dans cette étude, nous nous posons trois questions principales :

1. Y-a-t-il des futures enseignantes tchèques qui réalisent le contraste entre les voyelles e/ɛ, o/ɔ, et ø/œ ? Si oui, sur quels indices acoustiques principaux s'appuient ces contrastes ?

2. Quel est l'effet du contexte (nul ou p, t, k, R) sur la réalisation du contraste ? Existe-t-il des contextes facilitants (par rapport à la LM tchèque des locutrices) la réalisation des contrastes e/ε, o/ɔ, et ø/œ ?
3. Enfin, la performance en perception conditionne-t-elle la performance en production comme l'affirment de nombreux chercheurs ? Si le résultat montre que toutes celles qui produisent un contraste le perçoivent, alors l'affirmation est confirmée. Si à l'inverse certaines futures enseignantes produisent un contraste qu'elles ne perçoivent pas, le lien de causalité entre perception et production est remis en question.

Les valeurs moyennes formantiques des voyelles finales produites en contextes [p, t, k, R] sont indiquées dans des tableaux pour chaque locutrice séparément (calculées à partir de dix répétitions) en annexe du chapitre 14 (pages 92 - 102 des annexes) où l'on peut également trouver les triangles vocaliques F1/F2 et F2/F3 indiquant la position exacte des dix répliques de chaque voyelle isolée (pages 102 – 112).

14.2.2.1 Contraste e/ε

Si l'on se base sur le critère de la distance acoustique de F1 et/ou F2 supérieur à un écart type, les données font apparaître que quatre locutrices (T2, T8, T9 et T10) réalisent deux qualités distinctes e/ε au moins dans quelques-uns des contextes étudiés.

La Figure 138 montre la valeur moyenne des trois premiers formants des voyelles [e] (un rond) et [ε] (une croix) répétés dix fois par la locutrice T8 en isolation (traits horizontaux traversant l'image) et dans des logatomes CVCVCVC où C est la consonne [p, t, k, R]. Observons les formants des syllabes finales.

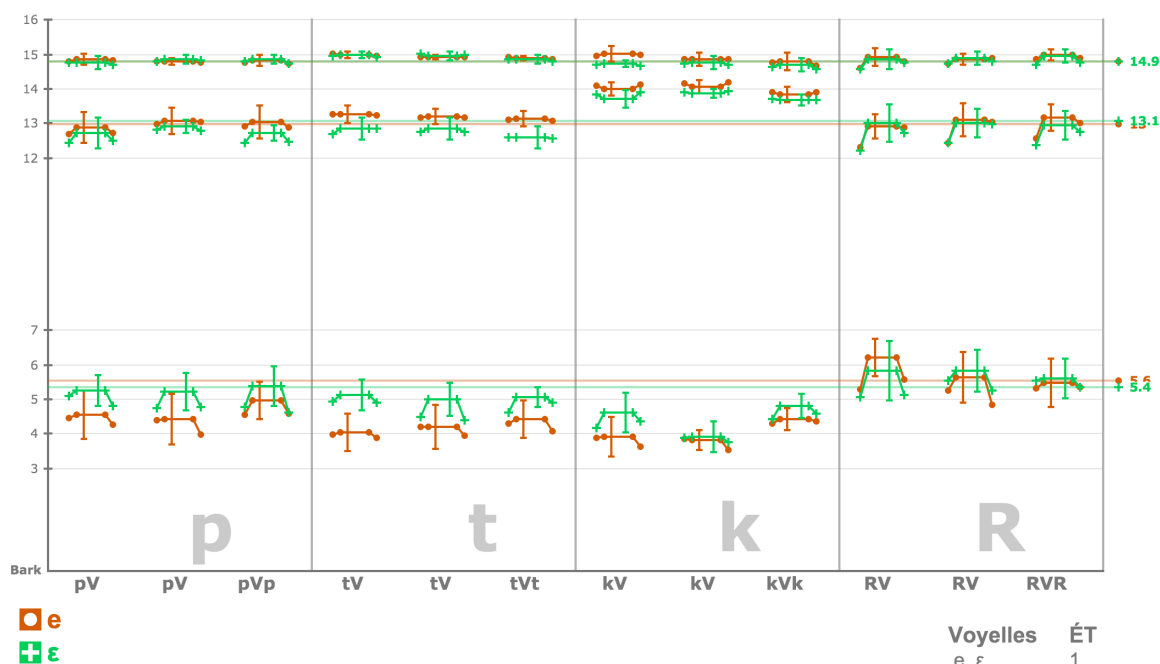


Figure 138 : Valeurs moyennes de F1, F2 et F3 (en Bark) de e/ε produit par la locutrice T8 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, R], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. Il est à noter la différence de valeurs de F1 et F2 en syllabe finale tVt

La Figure 138 montre que la locutrice T8 réalise le contraste entre e/ε est en syllabe finale tVt où les deux voyelles diffèrent par la valeur des premier et deuxième formants qui sont à plus d'un écart type l'un de l'autre : le [e] a alors un caractère plus diffus, avec F1 plus élevée et F2 plus bas que [ε] qui a un caractère plus compact.

Le triangle F1/F2 (en Bark) de la Figure 139 montre la position acoustique des dix répliques du [e] et du [ε] produites en syllabe finale tVt par la locutrice T8 : la majorité des sons représentant la qualité [e] sont produits avec F1 plus bas et F2 plus élevé que la majorité de sons représentant la qualité [ε].

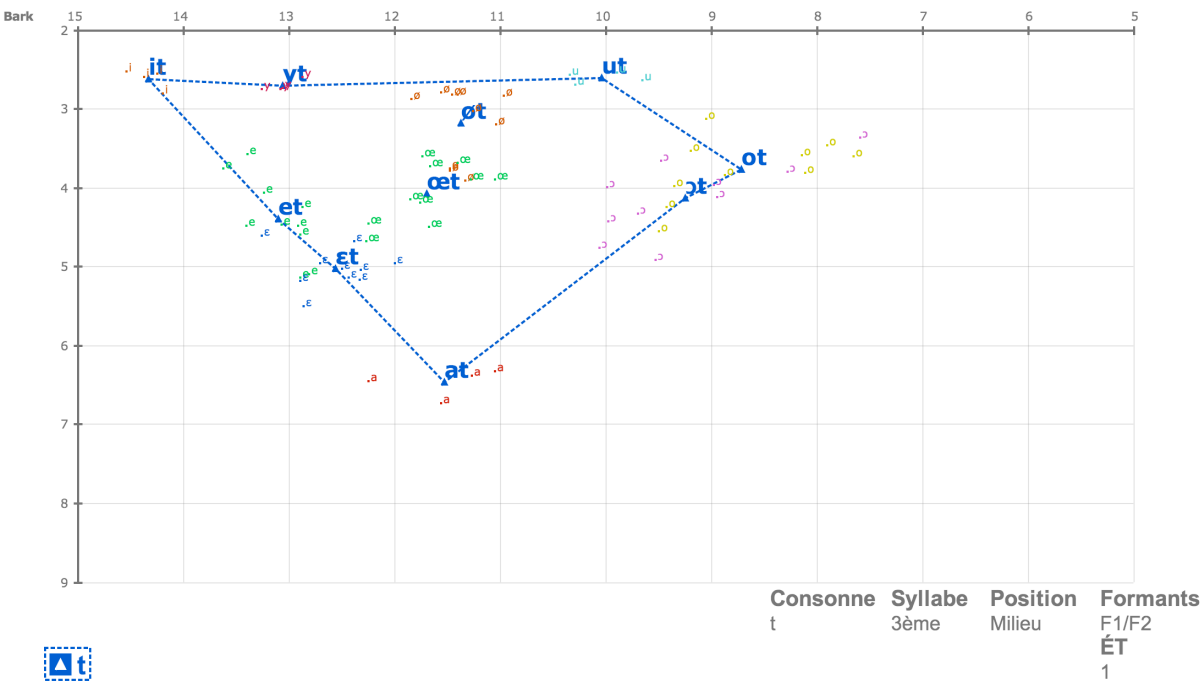


Figure 139 : Triangle F1/F2 (en Bark) des voyelles du français prononcées 10 fois par la locutrice T8 en syllabe finale tVt (valeur prise à la moitié de la durée vocalique). Il est à noter que e/ε sont pour la plupart séparés sur le plan F1 et F2

Le contraste entre e/ε est ensuite réalisé par la locutrice T9. Les trois premiers formants moyens de la voyelle [e] (un rond) et [ε] (une croix), calculés à partir de dix répétitions de la locutrice T9 sont illustrés à la Figure 140.

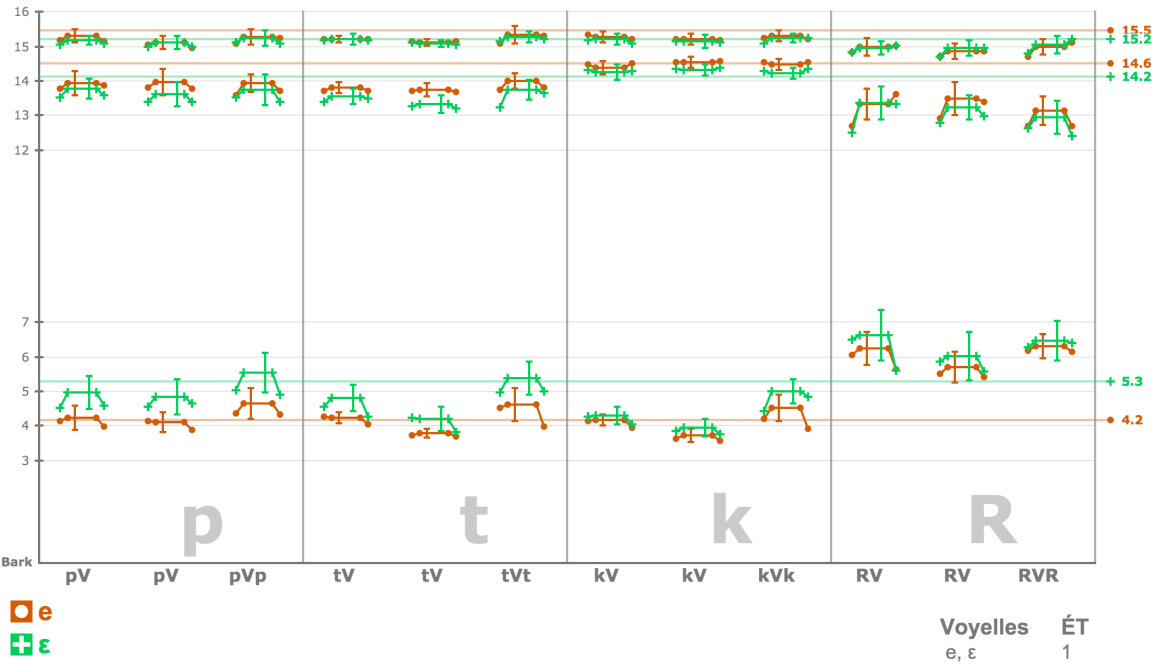


Figure 140 : Valeurs moyennes de F1, F2 et F3 (en Bark) de e/ε produit par la locutrice T9 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. Notons la différence de valeurs de F1 en isolation et en syllabe finale pVp, tVt et kVt

La Figure 140 montre que la locutrice T9 réalise le contraste e/ε en isolation et en syllabes finales pVp, tVt et kVk car leur premier formant se trouve à plus d'un écart type l'un de l'autre. Contrairement au contraste réalisé en contexte tVt par la locutrice T8, la valeur de F2 des deux voyelles est semblable. Les triangles vocaliques F1/F2 indiquant la position acoustique des dix répliques de [e] et [ε] produites par la locutrice T9 en isolation et en contextes [p, t, k] se trouvent commentés en annexe du chapitre 14, à la page 111 puis 149 et 150 des annexes. L'analyse des dix répliques montre que la locutrice T9 réalise le contraste dans la plupart des contextes phonétiques étudiés en séparant les deux qualités e/ε par la valeur de leur F1. Il reste à vérifier dans une étude ultérieure si le contraste est suffisamment saillant pour être perçu par les Français natifs.

Ensuite, la locutrice T10 produit également des sons représentant deux qualités acoustiquement distinctes e/ε. Les valeurs formantiques moyennes calculées à partir de valeurs relevées à un tiers, à la moitié et à deux tiers des voyelles [e] et [ε] répétées dix fois par la locutrice T10 en isolation (traits horizontaux) et dans des logatomes CVCVCVC où C correspond à la consonne [p, t, k, ʁ] sont illustrées à la Figure 141.

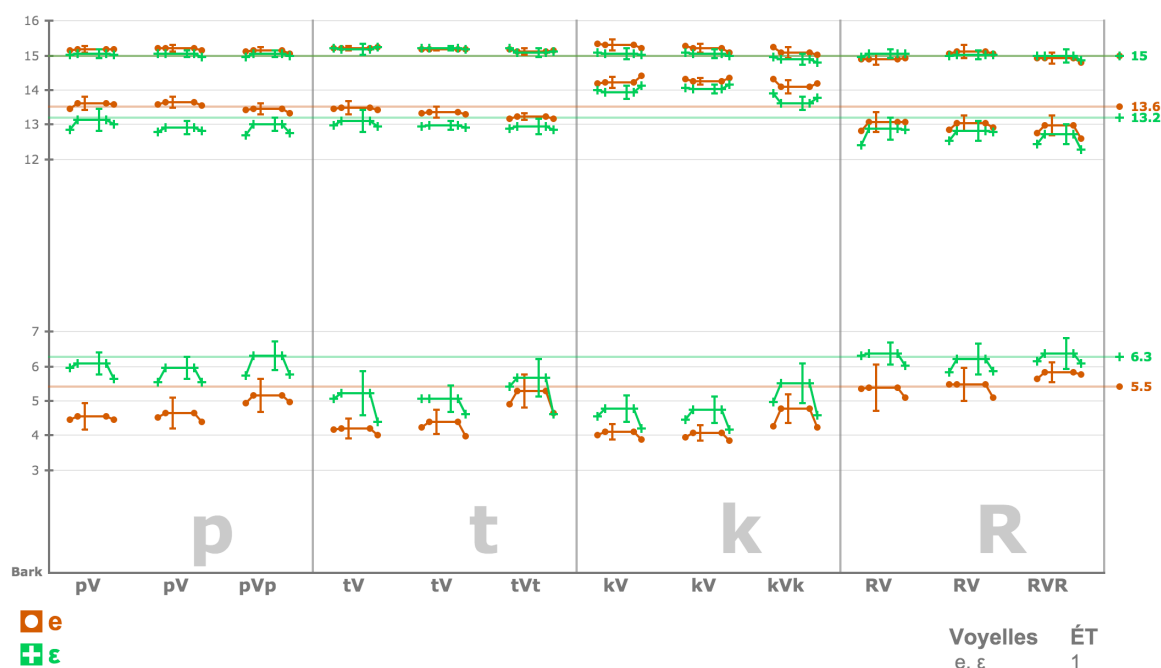


Figure 141 : Valeurs moyennes de F1, F2 et F3 (en Bark) de e/ε produit par la locutrice T10 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. Notons la différence de valeurs de F1 et/ou F2 en isolation et en syllabe finale pVp, kVk et RVR

La Figure 141 montre que la locutrice T10 réalise le contraste entre e/ε en isolation et en syllabes finales pVp, kVk et RVR. Alors qu'en contexte uvulaire le contraste est essentiellement marqué par la valeur du premier formant qui est à plus d'un écart type l'un de l'autre, en isolation, en contextes labial et palato-vélaire le contraste est marqué sur les plans F1 et F2. Les réalisations acoustiques de chacune des dix répliques de e/ε en isolation et en syllabes finales pVp, kVk et RVR sont illustrées dans les triangles F1/F2 (en Bark) qui se trouvent commentés en annexe du chapitre 14, à la page 112, puis 150 - 152 des annexes.

Enfin, le contraste e/ε est également réalisé par la locutrice T2 en syllabe finale pVp et tVt. Contrairement aux Français natifs dont les deux timbres s'opposent essentiellement par la valeur de F1, le contraste e/ε produit par la locutrice T2 est basé sur l'indice acoustique de F2. La Figure 107 de la page 260 illustre les valeurs moyennes des trois premiers formants de [e] (un rond) et [ε] (une croix) produits dix fois par la locutrice T2 en isolation (traits horizontaux) et dans des logatomes CVCVCVC où C représente la consonne [p, t, k, ʁ].

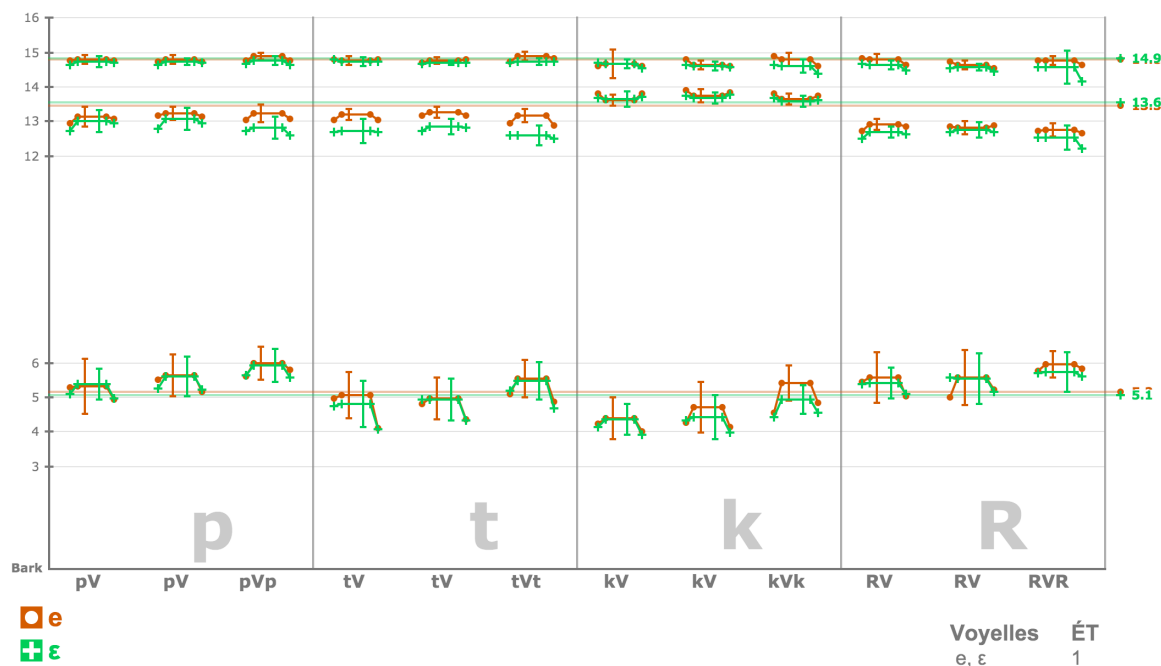


Figure 142 : Valeurs moyennes de F1, F2 et F3 (en Bark) de e/ε produit par la locutrice T2 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVCV où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. L'écart type tracé est de 1. Notons la différence de valeurs de F2 en syllabe finale pVp et tVt

Le triangle F1/F2 (en Bark) de la Figure 143 met en évidence que le contraste e/ε en syllabe finale pVp (valeur prise à la moitié de la durée vocalique) est pour la plupart des sons marqué par la locutrice T2 au niveau du deuxième formant, corrélé à l'antériorité-postériorité, avec [e] ayant F2 plus élevé (entre approximativement 2050 Ha et 2250 Hz) que [ε] (entre 1850 Hz et 2050 Hz). Notons qu'il ne s'agit pas de l'indice acoustique principal chez les natifs, comme nous l'avons montré en 13.2.2.1.

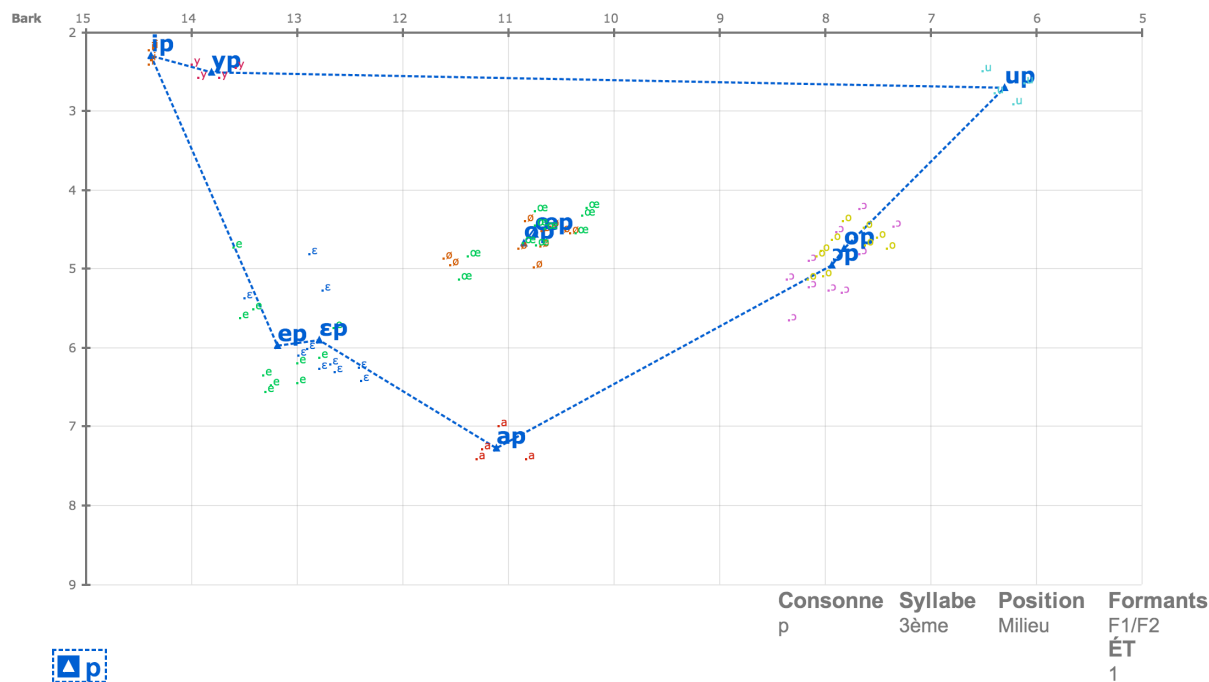


Figure 143 : Triangle F1/F2 (en Bark) des voyelles du français prononcées 10 fois par la locutrice T2 en syllabe finale pVp (valeur prise à la moitié de la durée vocalique). Notons que les occurrences des e /ε sont pour la plupart séparés sur le plan F2

Enfin le contraste marqué par la locutrice T2 est visible sur le plan F2 dans le cas de e/ε produits également en syllabe finale tVt, comme illustré à la Figure 144. La plupart des sons représentant la qualité [e] (avec F2 entre approximativement 2050 Hz et 2200 Hz) y sont clairement différenciés de ceux représentant la qualité [ε] (avec F2 souvent entre 1750 Hz et 1950 Hz).

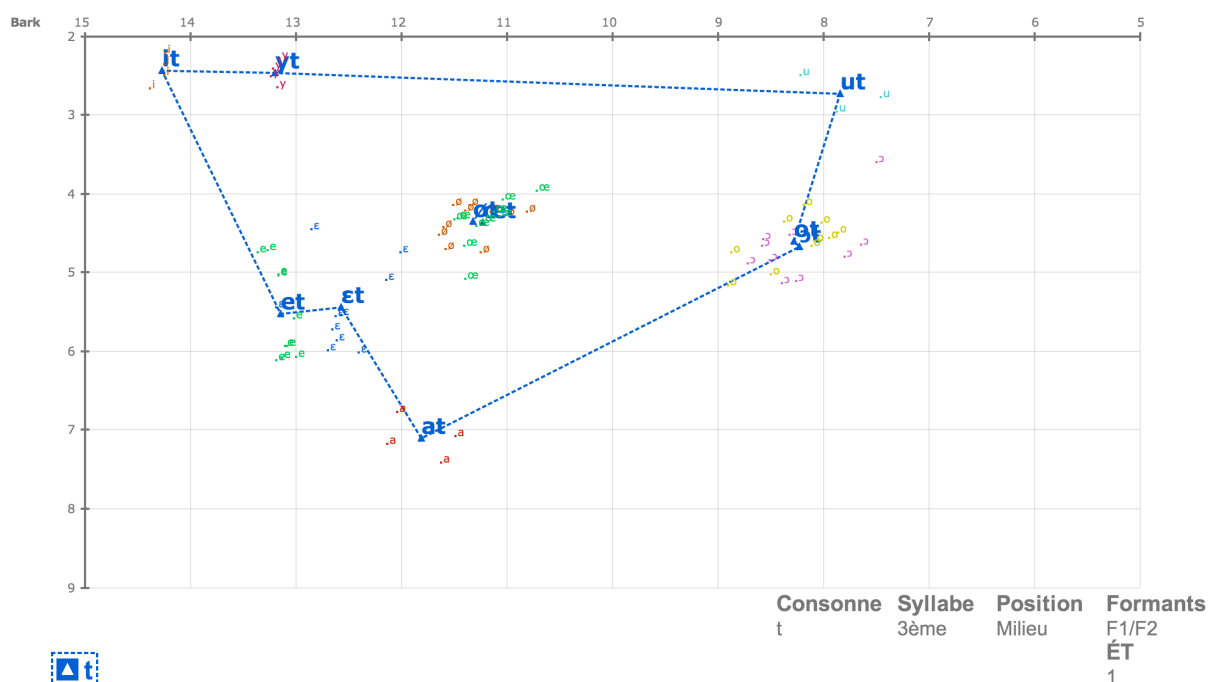


Figure 144 : Triangle F1/F2 (en Bark) des voyelles du français prononcées 10 fois par la locutrice T2 en syllabe finale tVt (valeur prise à la moitié de la durée vocalique). Notons que e (en vert)/ ε (en bleu) sont séparés sur le plan F2

Le contraste e/ε est ainsi réalisé par quatre futures enseignantes tchèques de FLE. Ces dernières produisent cependant des voyelles qui contrastent par la valeur de F1 dans une moindre mesure que la “référence” et le contraste de la locutrice T2 se base sur le seul indice acoustique de F2 (qui est un indice secondaire).

Pour résumer, trois locutrices sur quatre marquent le contraste en contextes labial et coronal qui semblent alors faciliter sa réalisation. En effet, dans ces contextes, [e] français est maximale-*ment* *similaire* – acoustiquement et perceptivement – au [ɪ] tchèque alors que [ε] français y est maximale-*ment* *similaire* aux [ε/ɛ:] tchèques. Ainsi, les deux timbres e/ε dans ces contextes sont assimilés à deux catégories maternelles distinctes dans l’esprit des tchécophones qui peuvent alors créer plus facilement deux catégories mentales pour ce contraste étranger.

En reliant ces résultats à ceux du test de discrimination de la section 14.1, nous remarquons que les locutrices T2, T8 et T9 qui réalisent les contrastes en production, les discriminent avec grande difficulté (dans seulement 2 suites sur 16). En revanche on observe un lien entre perception et production chez la locutrice T10 qui a discriminé le contraste e/ε dans 14 suites sur 16 et a produit ce même contraste dans la plupart des contextes consonantiques.

14.2.2.2 Contraste ø/œ

Le contraste ø/œ est produit par deux locutrices car elles réalisent ces voyelles avec des formants F1 et/ou F2 à plus d’un écart type les uns des autres.

Le contraste repose chez la locutrice T8 sur F1 en isolation et en syllabe finale tVt et sur F1 et F2 en syllabe finale pVp. La Figure 145 (page 336) montre les valeurs moyennes F1, F2 et F3 des voyelles ø/œ, calculées à partir de dix répétitions par la locutrice T8 en isolation (trait horizontal) et dans des logatomes CVCVCVC où C est la consonne [p, t, k, ʁ].

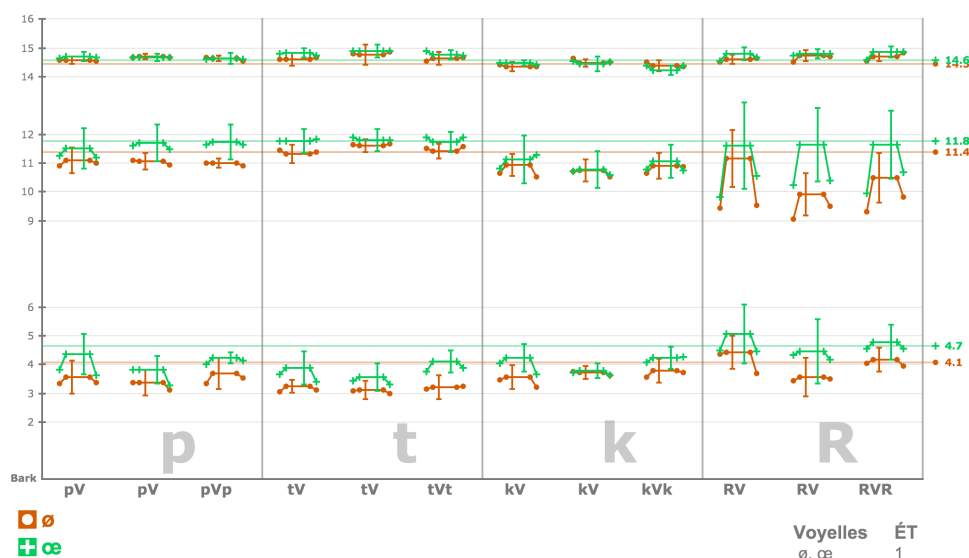


Figure 145 : Valeurs moyennes de F1, F2 et F3 (en Bark) de /ø/œ/ produit par la locutrice T8 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. L'écart type tracé est de 1. Notons la différence de valeurs de F1 et/ou F2 en isolation et en syllabe finale pVp et tVt

La position acoustique exacte de chacune des dix répliques de /ø/œ/ en isolation et en syllabe finale pVp et tVt réalisées par la locutrice T8 est illustrée dans les triangles vocaliques F1/F2 qui se trouvent en annexe du chapitre 14 (page 152). Les résultats montrent que même si la plupart des sons représentant les deux qualités sont clairement distingués par la valeur de F1 en contexte tVt, puis par les valeurs F1 et F2 en contexte pVp, certaines répliques se trouvent dans les triangles parmi les sons représentant l'autre catégorie. **Ce caractère fluctuant révèle que les voyelles ne sont pas tout à fait maîtrisées mais que le processus d'apprentissage est en cours.**

La deuxième future enseignante de FLE qui réalise le contraste entre /ø/œ/ est la locutrice T9. La Figure 146 illustre les valeurs moyennes de F1, F2 et F3 des voyelles [ø] et [œ] produites dix fois par la locutrice T9 en isolation (traits horizontaux) et dans des logatomes CVCVCVC où C représente l'une des consonnes [p, t, k, ʁ]. Les valeurs ont été relevées à un tiers, à la moitié et à deux tiers de la durée vocalique.

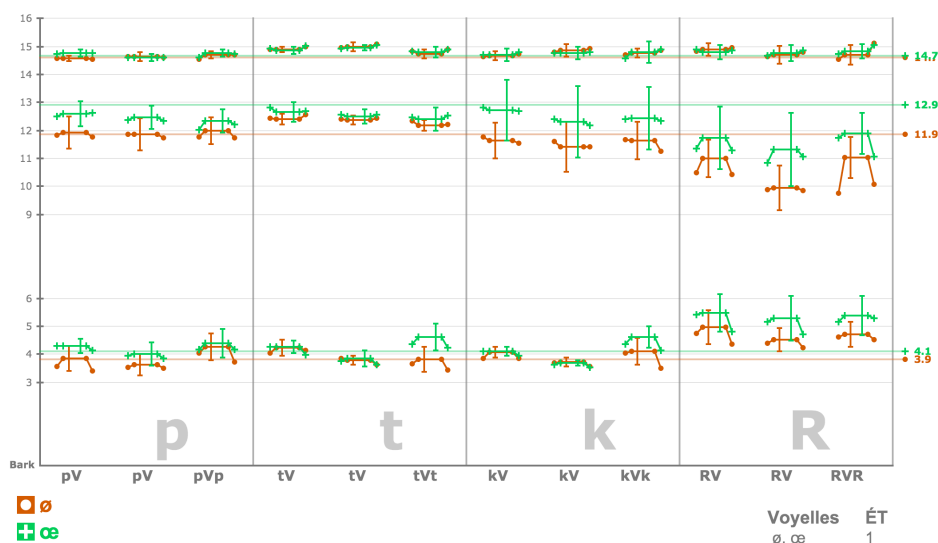


Figure 146 : Valeurs moyennes de F1, F2 et F3 (en Bark) de /ø/œ/ produit par la locutrice T9 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. L'écart type tracé est de 1. Notons la différence de valeurs de F1 et/ou F2 en isolation et en syllabe finale tVt

Le triangle F1/F2 de la page 109 des annexes montre qu'en isolation, la locutrice T9 produit deux qualités vocaliques qui diffèrent par la valeur de leur F2. Il ne s'agit pas cependant d'indice acoustique principal qui est à la base du contraste $\emptyset/\text{œ}$ chez les Français natifs. Les sons vocaliques représentant la qualité $[\emptyset]$ y sont alors réalisés avec F2 entre 1580 Hz et 1820 Hz alors que ceux représentant la qualité $[\text{œ}]$ avec F2 entre 1850 Hz et 2150 Hz.

En revanche, la locutrice T9 réalise le contraste $\emptyset/\text{œ}$ en syllabe finale tVt en qui est basé sur le même indice acoustique que la "référence", à savoir essentiellement le F1. En effet, les sons de la qualité visée $[\emptyset]$ se réalisent le plus souvent avec F1 entre 300 Hz et 400 Hz alors que ceux de la qualité visée $[\text{œ}]$ avec F1 le plus souvent entre 450 Hz et 550 Hz. Cette séparation ne concerne néanmoins pas tous les sons produits car trois répliques du $[\emptyset]$ ont été réalisées avec F1 entre 400 Hz et 500 Hz et une réplique du $[\text{œ}]$ avec F1 aux alentours de 350 Hz.

Quoique les locutrices T8 et T9 montrent du progrès dans la réalisation de l'opposition $\emptyset/\text{œ}$, elles discriminent ces deux voyelles avec difficulté (dans respectivement 4 et 5 suites sur 16). **Ainsi le rapport de cause à effet entre la perception et production du contraste $\emptyset/\text{œ}$ par les Tchèques T8 et T9 n'est pas confirmé.**

Pour résumer, le contraste entre $\emptyset/\text{œ}$ semble plus difficile à réaliser que celui entre $e/\text{ɛ}$ car seulement deux locutrices différencient les sons des deux catégories au niveau acoustique. Le contexte nul (isolation) et dental semblent faciliter la réalisation du contraste ce qui pourrait s'expliquer par les résultats de la deuxième partie (voir le Tableau 90 de la page 231) à propos de la similarité interlangue des voyelles. En effet, en isolation les voyelles $[\emptyset]$ et $[\text{œ}]$ sont acoustiquement *nouvelles* (elles se trouvent dans une partie de l'espace acoustique F1/F2 inoccupée par des voyelles du tchèque) ce qui semble faciliter la création de deux nouvelles catégories. Puis en contexte dental, $[\text{œ}]$ est perceptivement et acoustiquement *similaire* aux $[\text{ɛ}/\text{ɛ}:]$ tchèques alors que $[\emptyset]$ est dans ce contexte assimilé aux catégories tchèques $/u/u:/$. Le contraste renvoie donc à deux catégories maternelles différentes.

14.2.2.3 Contraste $o/\text{ɔ}$

Le contraste entre $o/\text{ɔ}$ est basé sur deux indices acoustiques principaux chez les Français natifs : le F1 et le F2. Il est donc nécessaire d'opposer les qualités en question par les deux formants afin que la différence acoustique soit perceptivement saillante.

Les futures enseignantes tchèques de FLE qui marquent le contraste entre $o/\text{ɔ}$ sont au nombre de trois et il s'agit de la locutrice T1, T2 et T10.

Les valeurs moyennes de F1 et F2 des voyelles $[o]$ (un rond) et $[\text{ɔ}]$ (une croix) répétées dix fois par la locutrice T10 en isolation (traits horizontaux) et dans des logatomes CVCVCVC avec C qui correspond à la consonne $[p, t, k, \text{ʁ}]$ se trouvent à la Figure 147.

La Figure 147 montre que le contraste entre $o/\text{ɔ}$ de la locutrice T10 se base en isolation et en syllabes finales pVp et RVR sur les formants F1 et F2, comme la "référence", et en syllabes finales tVt et kVk essentiellement sur le F2 (avec les valeurs de F1 à moins d'un écart type les unes des autres). Pour des raisons de fluidité de la lecture, les triangles vocaliques F1/F2 indiquant la position de chacune des dix répliques de $o/\text{ɔ}$ produits par la locutrice T10 dans les contextes mentionnés se trouvent en annexe du chapitre 14, pages 151 - 152 des annexes.

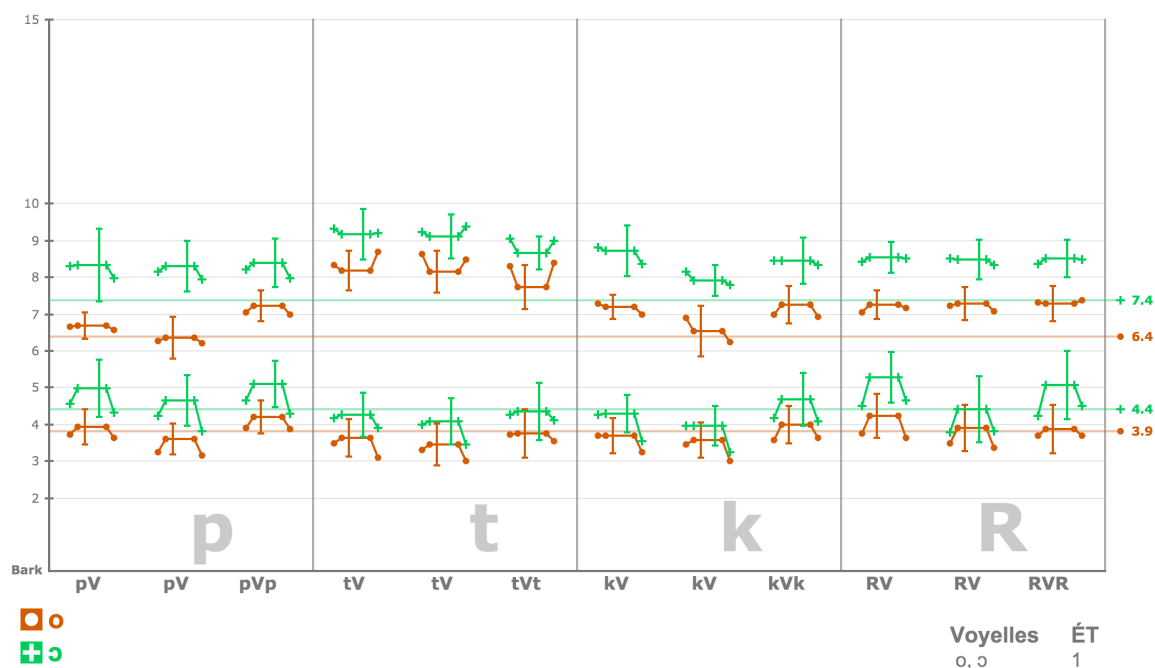


Figure 147 : Valeurs moyennes de F1, F2 et F3 (en Bark) de /o/ produit par la locutrice T10 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. L'écart type tracé est de 1. Notons la différence de valeurs de F1 et/ou F2 des voyelles isolées et dans des syllabes finales des quatre contextes

Le contraste o/ɔ est ensuite réalisé par la locutrice T1, comme on peut le lire à la Figure 148 qui montre les valeurs moyennes de F1, F2 et F3 de ces voyelles (calculées à partir de dix répliques) produites en isolation (traits horizontaux) et dans des logatomes CVCVCVC où C correspond à l'une des consonnes [p, t, k, ʁ].

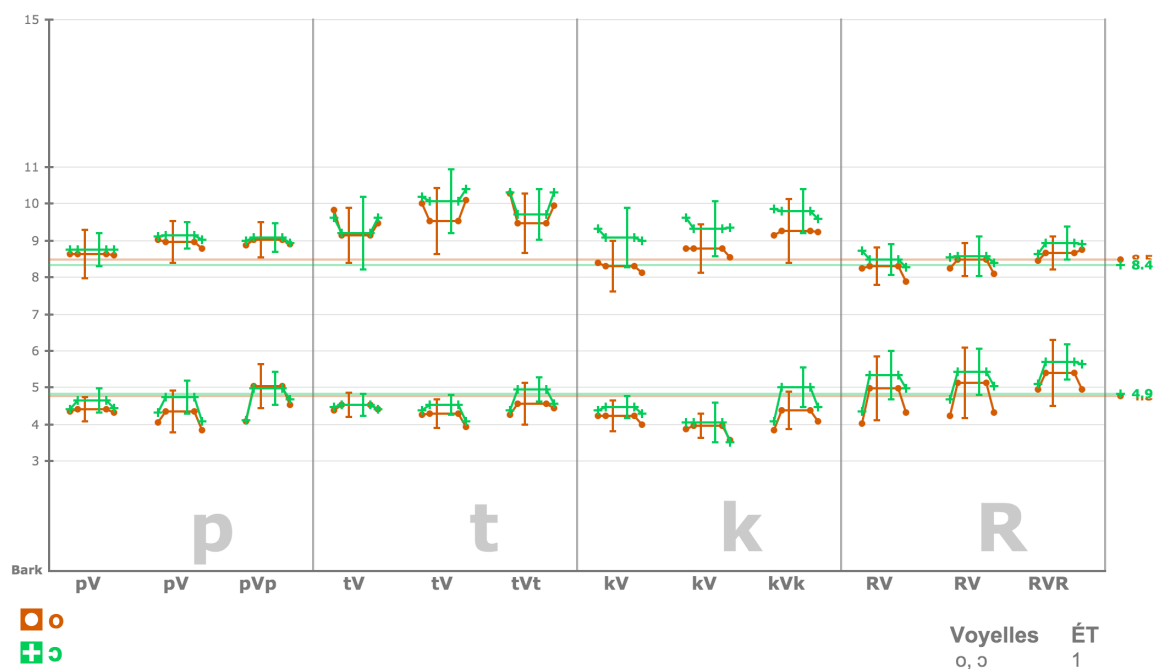


Figure 148 : Valeurs moyennes de F1, F2 et F3 (en Bark) de /o/ produit par la locutrice T1 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. L'écart type tracé est de 1. Notons la différence de valeurs de F1 en syllabe finale kVk

La Figure 148 montre que la locutrice T1 marque le contraste en syllabe finale kVk où le premier formant du [o] et du [ɔ] se trouve à plus d'un écart type l'un de l'autre. En isolation et dans les autres contextes, les voyelles sont réalisées avec des caractéristiques acoustiques semblables. La position acoustique des dix répliques de chacune des voyelles est illustrée dans le triangle vocalique F1/F2 des annexes (pages 152 - 153). Nous y remarquons que quatre répliques de [o] se trouvent dans la zone de réalisation de [ɔ] avec F1 et F2 élevés.

Enfin la locutrice T2 réalise également le contraste uniquement en syllabe finale kVk. De cette façon, la Figure 149 avec les valeurs moyennes de F1, F2 et F3 de [o] (un rond) et [ɔ] (une croix), calculées à partir de dix répliques produites en isolation (traits horizontaux) et dans des logatomes CVCVCVC où C correspond à l'une des consonnes [p, t, k, ʁ] montre que les deux voyelles varient en syllabe finale kVk par la valeur de leur premier formant qui se trouve à plus d'un écart type l'un de l'autre.

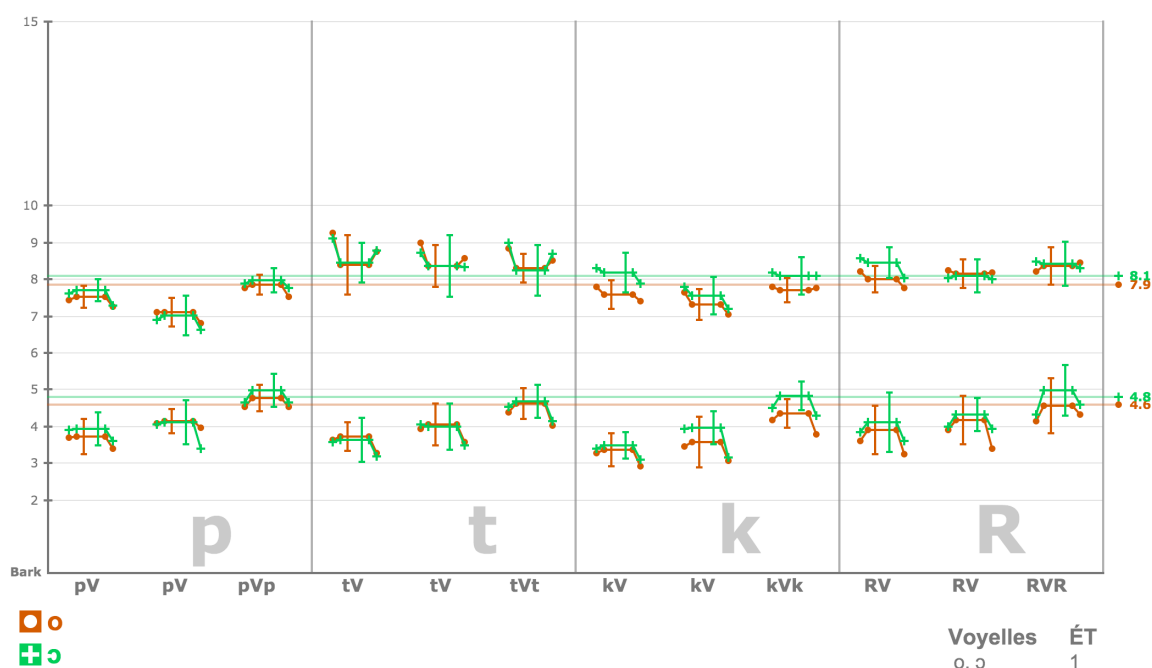


Figure 149 : Valeurs moyennes de F1, F2 et F3 (en Bark) de o/ɔ produit par la locutrice T2 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. Notons la différence de valeurs moyennes de F1 en syllabe finale kVk

La position acoustique exacte de chacune des dix répliques de [o] et [ɔ] de la locutrice T2 en syllabe finale kVk est indiquée dans le triangle vocalique F1/F2 en annexe du chapitre 14 (page 153 des annexes). Cette figure montre que le contraste n'est pas complètement acquis car trois répliques de [o] se réalisent avec F1 d'environ 500 Hz ce qui les place dans l'espace acoustique des sons correspondant au [ɔ].

Pour résumer, seulement une future enseignante (locutrice T10) réalise le contraste entre les voyelles [o] et [ɔ] en isolation et dans tous les contextes étudiés [p, t, k, ʁ] basé le plus souvent sur les indices acoustiques F1 et F2, comme la "référence". Deux autres futures enseignantes (locutrice T1 et T2) opposent la plupart des sons représentant les deux qualités vocaliques seulement en syllabe finale kVk et uniquement au plan F1. Le contexte palato-vélaire semble alors faciliter la réalisation du contraste ce qui peut s'expliquer de nouveau par la similarité interlangue des voyelles. Comme le révèle le résultat de la deuxième partie (voir le Tableau 90 de la page 231), [o] est dans ce contexte maximalement perceptivement *similaire* aux [u/u:] tchèques alors que [ɔ] aux [o/o:] tchèques. De cette façon, les deux timbres peuvent renvoyer aux deux catégories différentes dans l'esprit des tchécophones même si dans l'ensemble cette opposition s'avère difficile.

Concernant le rapport entre la perception et production du contraste o/ɔ, il n'est évident que chez la future enseignante T10 qui a discriminé ces deux voyelles dans 10 suites sur 16 et qui le réalise dans tous les contextes. Quoique les Tchèques T1 et T2 réalisent le contraste entre o/ɔ en contexte kVk, elles le discriminent avec difficulté (dans respectivement 2 et 4 suites sur 16). **Le résultat ne nous permet pas de confirmer l'affirmation selon laquelle la performance en perception conditionne celle en production.**

14.3 Voyelles moyennes chez des futures enseignantes tchèques : Conclusion partielle

Les résultats de l'étude qualitative du chapitre douze concernant la perception et la production des voyelles moyennes par des futures enseignantes de FLE de LM tchèque doivent être considérés comme un complément aux résultats de l'étude quantitative du chapitre 12 qui décrit l'approximation possible de la "référence" par tout le groupe des tchécophones et non pas comme une contradiction. Ils montrent en effet qu'en étudiant les productions de différentes futures enseignantes séparément, certaines atteignent l'authenticité qui n'est pas attendue sur la base du SLM. Il s'agit notamment de l'acquisition authentique des voyelles *similaires* [e, o] par une tchécophone, dont l'atteinte authentique n'est pas attendue.

Le chapitre 14 révèle également que pour déterminer le niveau phonético-phonologique en L2, il n'est pas possible de se fier uniquement à l'évaluation perceptive car l'oreille humaine est incapable de noter certains progrès pourtant tangibles quand on regarde les résultats acoustiques. La comparaison des formants au sein de chaque paire des voyelles montre en effet que certaines futures enseignantes réalisent, du moins dans certains contextes, les contrastes entre les voyelles moyennes et que le processus d'apprentissage est donc en cours. Les contrastes entre les voyelles moyennes réalisés en isolation par les futures enseignantes de FLE ne sont cependant pas toujours perçus car ils se basent soit sur des indices acoustiques différents de ceux de la "référence", soit sur le bon indice mais la différence acoustique entre les deux voyelles n'est pas assez grande.

L'étude montre que l'opposition e/ɛ est la plus souvent réalisée par les futures enseignantes, du moins en contextes labial et dental. Dans ces contextes, les deux qualités françaises ont une ressemblance acoustique et perceptive maximale avec deux qualités tchèques distinctes (respectivement [ɪ] et [ɛ/ɛ:]) ce qui pourrait expliquer pourquoi la création de deux catégories mentales e/ɛ est plus facile dans ces contextes.

La réalisation des contrastes o/ɔ et ø/œ est moins courante et dépend également du contexte. Alors que le contraste o/ɔ est le plus souvent réalisé en contexte palato-vélaire où les deux voyelles sont assimilées à deux catégories tchèques différentes (respectivement aux /u/u/ et /o/o:/), le contraste ø/œ est le plus souvent réalisé en contexte nul où les voyelles sont acoustiquement *nouvelles*, puis en contexte dental où [ø] renvoie aux /u/u:/ tchèques et [œ] aux /ɛ/ɛ:/ tchèques. Ainsi il existe des contextes facilitants l'apprentissage des contrastes qui sont définis dans ce travail par rapport à la LM de l'apprenant et non pas par rapport à l'entourage consonantique (contexte assombrissant/ éclaircissant) comme dans la méthode verbo-tonale (Guberina *et al.*, 1965).

A la fin de ce chapitre, nous encourageons donc les futures enseignantes à prendre conscience de l'existence de six qualités vocaliques différentes [e, ɛ, o, ɔ, ø, œ] et de renforcer le contraste entre les deux membres de chaque paire qui dans le cas de e/ɛ et ø/œ est essentiellement basé sur l'indice acoustique de F1 (corrélé à l'aperture) et dans le cas de o/ɔ sur F1 et F2 (corrélé à l'antériorité-postériorité et arrondissement).

Conclusion finale et discussions

La bêtise des gens consiste à avoir une réponse à tout. La sagesse consiste à avoir une question à tout.

Milan Kundera

Nous nous sommes intéressée dans cette thèse aux limites de la maîtrise phonétique des dix voyelles orales du français [i, e, ε, a, u, o, ɔ, y, ø, œ] par dix futures enseignantes de FLE de langue maternelle tchèque, censées posséder une performance quasi-native de la langue enseignée.

Nous avons orienté notre intérêt vers l'acquisition des voyelles en isolation (qui n'apparaissent dans la parole naturelle que très rarement), et en contexte (contextes consonantiques et positions prosodiques divers). En effet seulement quand tous les allophones d'une voyelle sont acquis, l'on peut estimer que la catégorie phonémique est maîtrisée (Flege, 1991).

Afin de classer les voyelles du français selon la possibilité de leur acquisition authentique par les futures enseignantes tchèques de FLE, nous avons utilisé le modèle d'apprentissage phonétique SLM (Speech Learning Model) de Flege dont les hypothèses et postulats sont clairement exposés par l'auteur (Flege, 1995). La préférence pour l'application de ce modèle particulier par rapport aux autres modèles très prisés, notamment le Perceptual Assimilation Model (PAM) décrit par Best (1995) prête à discussion. En effet, selon certains chercheurs, ce dernier est plus fiable que le SLM car il permet de prédire les écarts effectifs des apprenants avec une plus grande précision (Best *et al.*, 2003).

Pour expliquer notre choix, il est important de rappeler les fondements de ces deux modèles. PAM est sans aucun doute un modèle puissant dans la prédiction des difficultés qu'auront les apprenants débutants lors de l'identification des sons et notamment de discrimination de contrastes des sons étrangers. Ses hypothèses telles qu'exposées en 1995, sont construites sur la base d'études de la **perception** des auditeurs qui ne connaissent pas encore la LE (que l'on peut alors considérer comme apprenants **grands débutants**). Il est évident qu'au stade initial de l'apprentissage, la perception des sons étrangers par des apprenants ayant la même langue maternelle (LM) est influencée uniquement par cette dernière (qui est donc le **seul facteur** affectant la perception des auditeurs naïfs). Le groupe d'apprenants visé explique pourquoi les prédictions de PAM concernant les difficultés sont alors souvent fiables et pourquoi elles s'appliquent à tous les apprenants du même arrière-plan linguistique de la même façon. Ainsi PAM est un modèle purement perceptif qui sous-entend que tous les écarts en production s'expliquent par la perception erronée de l'apprenant. Or comme nous l'avons montré dans le chapitre 14, le rapport de causalité entre perception et production est loin d'être évident.

A l'inverse, SLM fournit des hypothèses concernant le niveau phonétique des apprenants **avancés** ayant appris la LE tardivement. Ces derniers peuvent maîtriser fidèlement les voyelles de la LE *identiques* à celles de la LM, acquérir avec authenticité les voyelles *nouvelles* et de produire et percevoir les voyelles de la LE *similaires* à celles de la LM par *mécanisme de classification par équivalence* (décrit également comme *old wine in new bottle*) qui exclue l'authenticité. Les apprenants visés par le modèle ont derrière eux une longue expérience avec la LE et leur niveau phonétique effectif dépend donc de **nombreux facteurs**, tels que la qualité et quantité d'input, la motivation, etc. Ensuite, les hypothèses de SLM sont construites à partir de résultats d'études non seulement en **perception** mais également en **production**. Il s'agit donc d'un modèle relativement complexe quoique orienté uniquement vers les écarts au niveau segmental.

Pour définir quelles voyelles du français sont *nouvelles*, *identiques* et *similaires* aux voyelles tchèques, nous nous sommes appuyée sur trois critères de comparaison : 1. les symboles phonétiques, 2. les caractéristiques acoustiques, 3. les propriétés perceptives.

1. Les symboles phonétiques

Alors que la description phonologique des voyelles orales du français est constante et se fait par les symboles de l'API /i, e, ε, a, u, o, ɔ, y, ø, œ/, celle du système tchèque varie selon les auteurs. En phonologie du tchèque, où l'on décrit la fonction des phonèmes au sein du système, cinq symboles identiques représentant cinq timbres avec une différence de durée /i, i :, e, e :, o, o :, u, u :, a, a :/ sont traditionnellement utilisés (Ludvikova and Kraus, 1966; Mazlova, 1946). Puisque nous nous intéressons dans ce travail à la nature des voyelles tchèques (que nous comparons à la nature des voyelles du français), nous utilisons les symboles phonétiques employés dans la littérature récente (Dankovicova, 1997a) et il s'agit de [a, u, ɪ, ε, o] pour marquer le timbre le plus courant des voyelles brèves et [a :, u :, i :, ε :, o :] pour marquer le timbre le plus courant des voyelles longues. Nous montrons par ce choix des symboles phonétiques en tchèque que :

a) Le contraste entre les voyelles hautes [ɪ, i:] est basé non seulement sur la durée mais également sur le trait tendu/relâché comme c'est le cas de 20 % de langues avec des contrastes vocaliques de durée, répertoriées dans la base de données UPSID (Crothers, 1978). Ces symboles sont systématiquement utilisés dans la littérature récente.

b) Le contraste entre les voyelles hautes [u, u:] est essentiellement basé sur la durée selon nos données acoustiques et perceptives. Cependant Skarnitzl and Volin (2012), suite aux résultats d'une étude acoustique notent l'évolution du timbre de la voyelle brève vers [ʊ] qui serait plus relâché que [u:]. Le symbole [ʊ] est également utilisé par Dubeda (2005) ou Dubeda and Januska (2006).

c) Le symbole [o] représente mieux la qualité plutôt mi-fermée de la voyelle moyenne dans la plupart des contextes étudiés alors que le symbole [ε] reflète bien la qualité vocalique mi-ouverte dans tous les contextes étudiés. Notons que les auteurs utilisent parfois le symbole [ɔ] (Dubeda and Januska, 2006; Podlipsky, 2009).

L'inconstance dans le choix des symboles utilisés pour décrire le système vocalique tchèque remet en question la pertinence du critère de l'API dans la répartition des voyelles du français en *nouvelles*, *similaires* ou *identiques* aux voyelles tchèques. La faiblesse de ce critère a déjà été discutée par Rochet (1995) et la pertinence de la description des timbres vocaliques précis sur la seule base des symboles phonétiques a été remise en question par Vaissière (2007), (2009), (2011) : un même symbole représente souvent des timbres vocaliques différents.

En effet, selon la transcription que nous avons adoptée pour les voyelles tchèques, les voyelles du français « symboliquement » *nouvelles* pour les tchéco-phones seraient le [y, ø, œ, ɔ, e]. Mais si l'on se base sur les transcriptions d'autres chercheurs, les voyelles du français retranscrites avec [ɔ, e] sont *identiques* aux voyelles du tchèque. En considérant toutes les transcriptions possibles des voyelles tchèques que l'on trouve dans la littérature, seules les voyelles françaises notées par les symboles [y, ø, œ] sont « symboliquement » *nouvelles*. Les autres voyelles, à savoir [i, e, ε, a, u, o, ɔ] sont en revanche « symboliquement » *identiques*.

2. Les caractéristiques acoustiques

Puisque la comparaison des voyelles sur la seule base de leurs symboles phonétiques est clairement insuffisante, comme nous venons de l'illustrer, nous avons complété la comparaison des voyelles du tchèque et du français par l'étude de leurs propriétés acoustiques et perceptives. La similarité phonétique entre les voyelles est alors systématiquement étudiée en isolation et en liaison avec leur entourage consonantique (symétrique CVCVCVCV où C est [p, t, k]).

L'étude seule des voyelles isolées est insuffisante: il s'agit essentiellement d'un produit de laboratoire (Meunier, 2001). En revanche, les valeurs formantiques des voyelles isolées sont largement utilisées dans l'étude de la réduction phonétique car si l'on en croit Miller (1981), elles définissent la

corpus peut de nouveau prêter à discussion. D'un côté, les non mots sont souvent hyperarticulés (Durand, 1985) et les contextes consonantiques symétriques utilisés pour étudier les effets de la coarticulation ne sont pas représentatifs de la parole naturelle (Gendrot and Adda-Decker, 2010). De l'autre côté, un corpus constitué de logatomes permet de ne faire varier qu'un facteur à la fois et de dégager ainsi des régularités à propos des effets de la coarticulation ou de la position prosodique indépendamment des autres variations intervenant lors de la parole naturelle.

L'étude de la coarticulation par la comparaison des valeurs formantiques des voyelles en contexte par rapport à leur cible a montré que le formant F2 est maximalelement sensible au changement de l'entourage consonantique (mis à part le [i]), ce qui correspond aux résultats des études antérieures (Hillenbrand *et al.*, 2001; Stevens and House, 1963). En français et en tchèque, ces déplacements essentiellement horizontaux dans l'espace acoustique F1/F2 s'expliquent majoritairement par les loci des consonnes subséquentes et parfois par la centralisation acoustique. Ainsi le sens du mouvement de F2 est souvent constant dans les deux langues et correspond aux prédictions établies à partir de la synthèse articuloire : [p] ayant un locus bas attire en général les formants des voyelles environnantes vers le bas, sauf pour les voyelles postérieures du français. [t] avec un locus aux alentours de 1800 Hz (Delattre 1963) abaisse le et les formants reliés à la cavité antérieure des voyelles antérieures étirées et élève ceux des voyelles antérieures arrondies (présentes en français) et des voyelles postérieures. Enfin [k] élève le F2 et abaisse le F3 des voyelles proportionnellement à leur degré d'ouverture, avec les F2 et F3 des voyelles moyennes étant maximalelement affecté (créant ainsi « la pince vélaire ») et les F2 et F3 des voyelles fermées stables (sans mouvements importants).

Les deux langues diffèrent ensuite par l'ampleur des effets de coarticulation : nous avons montré une plus grande variation des voyelles antérieures tchèques par rapport aux voyelles antérieures du français et une plus grande variation de voyelles postérieures du français par rapport aux voyelles postérieures du tchèque. En effet puisque la coarticulation affecte maximalelement la valeur du formant (souvent F2) relié à la cavité antérieure, la variation de ce dernier par des Français natifs est limitée dans les hautes fréquences sans doute car il faut maintenir l'opposition entre les voyelles étirées (réalisées avec des valeurs de F2 plus élevées) et arrondies (réalisées avec des valeurs de F2 abaissées). En revanche les Tchèques, n'ayant qu'une série de voyelles antérieures dans leur système, ont une plus grande liberté dans les variations du F2. Ce résultat est attendu par Manuel and Krakow (1984) prédisant que les langues ayant un grand nombre de qualités vocaliques, comme le français, rencontreront des variations vocaliques moindres (afin de préserver les contrastes) que les langues avec un inventaire vocalique plus restreint, comme le tchèque. Ainsi, la baisse de F2 des voyelles antérieures brèves du tchèque (et notamment du [ɛ], engendrée par le contexte labial et dental) ne menace pas la neutralisation des contrastes avec les homologues arrondis (qui n'existent pas dans cette langue), de même que l'élévation de F2 des voyelles postérieures du français, notamment en contexte dental, ne menace pas la neutralisation de contrastes avec les homologues non arrondies (inexistants en français).

L'étude de la réduction phonétique suivant la position prosodique a ensuite mis en évidence un allongement final prépausal réalisé dans les logatomes trisyllabiques des deux langues, avec un rapport de 1 : 1 : 1,4 ou 1,5 (pour les voyelles brèves du tchèque). Pour confirmer que l'allongement de toute la syllabe finale CVC est de la même ampleur dans les deux langues (car l'allongement vocalique peut être compensé par un abrégement de la tenue consonantique, comme démontré par Adu Manyah and Sock (2002), il serait intéressant de mesurer dans le futur la durée de tous les segments constituant la syllabe finale.

Les variations spectrales en fonction de la position sont petites mais significatives et concernent en particulier le premier formant. Les mouvements sont alors essentiellement verticaux dans les triangles F1/F2 avec l'abaissement de F1 (inférieur à 70 Hz) notamment des voyelles ouvertes et mi-ouvertes en position médiane CVCVCVC (et initiale en tchèque CVCVCVC) par rapport à la position finale CVCVCVC où la voyelle est la plus longue.

Les contrastes entre les voyelles moyennes du français sont maintenus dans toutes les positions des logatomes, avec F1 des voyelles mi-fermées d'au moins 150 Hz plus bas que celui des voyelles mi-ouvertes. Ce résultat semble contredire les études sur le français (Delattre, 1969; Wioland, 2005) mais peut s'expliquer de nouveau par le choix du corpus basé sur des logatomes, souvent hyperarticulés (Durand, 1985). Pour cette raison, comme le précise van Bergem (1993, p. 21) : « *Stressing and destressing are properties that can only be assigned to existing words : it is difficult to simulate the natural destressing of syllables in nonsense words.* ». Ainsi, alors que notre corpus est adapté à l'étude de la coarticulation et permet de rendre compte de l'allongement prépausal, il s'avère moins pertinent dans l'étude des variations formantiques suivant la position par rapport à l'accent. De cette façon, dans la construction du corpus servant à rendre compte des réductions phonétiques dues à la position prosodique, il faudrait utiliser de vrais mots et veiller à ne faire varier qu'un paramètre à la fois, afin de ne pas mélanger l'effet de la position avec les effets de coarticulation, comme l'a fait Delattre (1969) (son corpus est précisé et analysé dans le Tableau 83 de la page 216).

Enfin, un dernier point concernant la réduction phonétique, observée et définie par la comparaison des formants avec la cible, prête à discussions. En français, le sens ou le degré de certaines variations formantiques des voyelles [a] et [ɔ] ne correspondent pas aux prédictions et à la littérature.

Selon nos données, dans la coarticulation des voyelles avec des consonnes, les mouvements de F1 par rapport à la cible sont en général minimes ne dépassant pas 50 Hz (sauf pour le contexte uvulaire) mais lorsque [ɔ] est prononcé en contexte labial, son F1 (corrélé à l'aperture) est de 57 Hz plus élevé que la cible (selon nos résultats, le F1 moyen de [ɔ] isolé est de 577 Hz). Ce phénomène ne s'explique ni par le locus de la consonne labiale qui attirerait le formant vers le bas, ni par une centralisation acoustique qui provoquerait de nouveau l'abaissement de F1 du [ɔ]. Notons au passage que le contexte labial provoque également l'élévation du deuxième formant de toutes les voyelles postérieures et du [a] français ce qui s'explique cette fois-ci par la centralisation acoustique qui élève le F2.

Ensuite, la très grande ampleur de la variation acoustique du [a] en contexte par rapport à la cible est un résultat frappant : le F2 de [a] s'élève par rapport à la cible de 133 Hz, 278 Hz, 515 Hz et 683 Hz en contextes respectivement [ʁ, p, t, k]. Pour comparaison, celui du [a] tchèque s'élève en contextes [t, k] respectivement de 121 Hz et 184 Hz et baisse en contexte [p] de 34 Hz.

Le principe selon lequel les voyelles isolées définissent la cible acoustique (définition de Miller (1981) mentionnée ci-dessus), doit donc être remis en question à cause des [a] et [ɔ] français. Comme il a déjà été noté par Vaissière dans ses cours de phonétique expérimentale de Master, le /a/ en isolation a souvent une réalisation pharyngale [ɑ] avec F1 élevé et F2 bas (selon nos résultats le F1 est en moyenne de 805 Hz et F2 de 1301 Hz). En revanche, une fois inséré en contexte consonantique et cela même en contexte uvulaire qui abaisse généralement le F2, son deuxième formant augmente au point qu'en contextes dental et palato-vélaire, [a] se retrouve dans le triangle vocalique dans la série de voyelles antérieures (avec F2 supérieur à 1800 Hz, comme illustré à la Figure 151).

Ensuite, le [ɔ] isolé n'apparaît jamais en syllabe finale ouverte dans la parole naturelle, comme nous l'avons rappelé ci-dessus. Sa production est donc nécessairement artificielle ce qui explique pourquoi certaines de ses variations acoustiques ne peuvent pas être dégagées à partir des formants de [ɔ] en isolation. A l'avenir, nous éviterons donc d'étudier la coarticulation de [a] et [ɔ] à partir de la position isolée et analyserons les mouvements relatifs induits par les différents contextes en distinguant entre l'effet centralisant et périphéralisant, comme l'ont fait Gendrot and Adda-Decker (2010) pour le français.

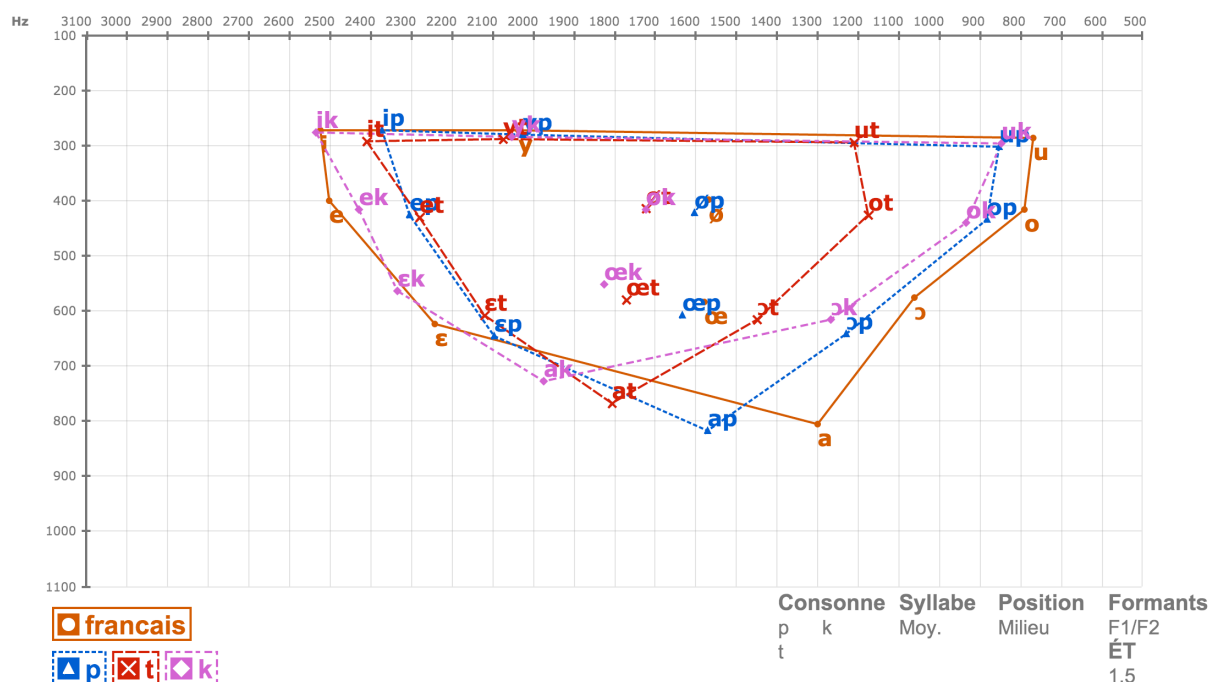


Figure 151 : Valeurs moyennes F1/F2 (en Bark) des voyelles orales du français produites dans des logatomes CVCVCVC où C = /p, t, k/, calculées à partir de 4 répétitions par dix Françaises (valeurs prises à la moitié vocalique)

3. Les propriétés perceptives

En comparant la similarité phonétique des voyelles du français aux voyelles du tchèque, les résultats mettent en évidence que cette dernière varie selon le contexte. Par exemple, le [a] français est perceptivement assimilé aux /a/a:/ tchèques en isolation et en contexte labial, et aux /ε/ε:/ tchèques en contextes dental et palato-vélaire. L'assimilation interlangue d'une même voyelle française à deux catégories maternelles tchèques distinctes en fonction du contexte s'explique par les données acoustiques commentées ci-dessus. Un autre exemple est le [o] français, acoustiquement maximale *similaire* au [u] tchèque en isolation et aux [o/o:] tchèques en contextes [p, t, k] (dont le F1 est dans ces contextes plus bas qu'en isolation). Les voyelles du français véritablement *nouvelles* (au niveau acoustique, perceptif et "symbolique") sont au nombre de deux et seulement dans certains contextes : [ø] est une voyelle *nouvelle* en isolation et en contexte palato-vélaire et [y] en contexte dental. Dans les autres contextes, ces voyelles sont perceptivement *similaires* à des voyelles tchèques.

Enfin les voyelles du français *identiques* aux voyelles du tchèque n'existent pas car le résultat du test-t indique qu'au moins un des formants comparés est différent.

Ces résultats ont des implications pour les modèles d'apprentissage phonétique basés sur la notion de similarité car ils mettent en évidence que :

- Sur la base de tests statistiques, la délimitation des sons en *identiques* (très rares) et *similaires* est presque impossible. Pour contourner ce problème, les auteurs utilisent parfois la notion de « hautement similaire » à la place de *identique* (Bohn and Flege, 1990; Strange *et al.*, 2005), que nous avons également adoptée.
- Tous** les sons vocaliques d'un système phonétique d'une langue (ici les voyelles du français) peuvent être au moins dans quelques contextes phonétiquement *similaires* aux voyelles d'un autre système (ici aux voyelles du tchèque) et selon le contexte, **la même** voyelle peut être *nouvelle* ou *similaire*.
- Lorsqu'on utilise la notion de similarité pour prédire les difficultés, il est donc important de prendre en compte le contexte phonémique. Mais que fait-on des contextes consonantiques qui

n'existent que dans la LE ? En effet l'apprentissage des voyelles du français en contexte uvulaire, qui n'existe pas en tchèque, ne peut pas être prédit sur la base de la similarité.

Notons que ces implications théoriques ont déjà été soulevées dans de nombreuses études (Gottfried, 1984; Levy, 2009; Levy and Strange, 2008; Strange *et al.*, 2001; Strange *et al.*, 2005).

Dans la troisième partie de la thèse, nous nous sommes intéressée à **l'approximation de la « référence » dans la production et perception** des dix voyelles orales du français par des futures enseignantes tchèques de FLE. Le résultat montre que les voyelles dont le F1 atteint des valeurs extrêmes, c'est-à-dire [i, y, u, a] sont mieux maîtrisées que les voyelles moyennes [o, ɔ, ø, œ, e, ε]. Comme nous pouvons le voir à la Figure 152, les voyelles [i, u, a, o, ɔ, e, ε] sont réalisées avec des caractéristiques acoustiques des voyelles maternelles (avec F1 et F2 à un écart type d'une voyelle tchèque ; valeurs prises à la moitié de la durée vocalique, tous contextes et syllabes confondus).

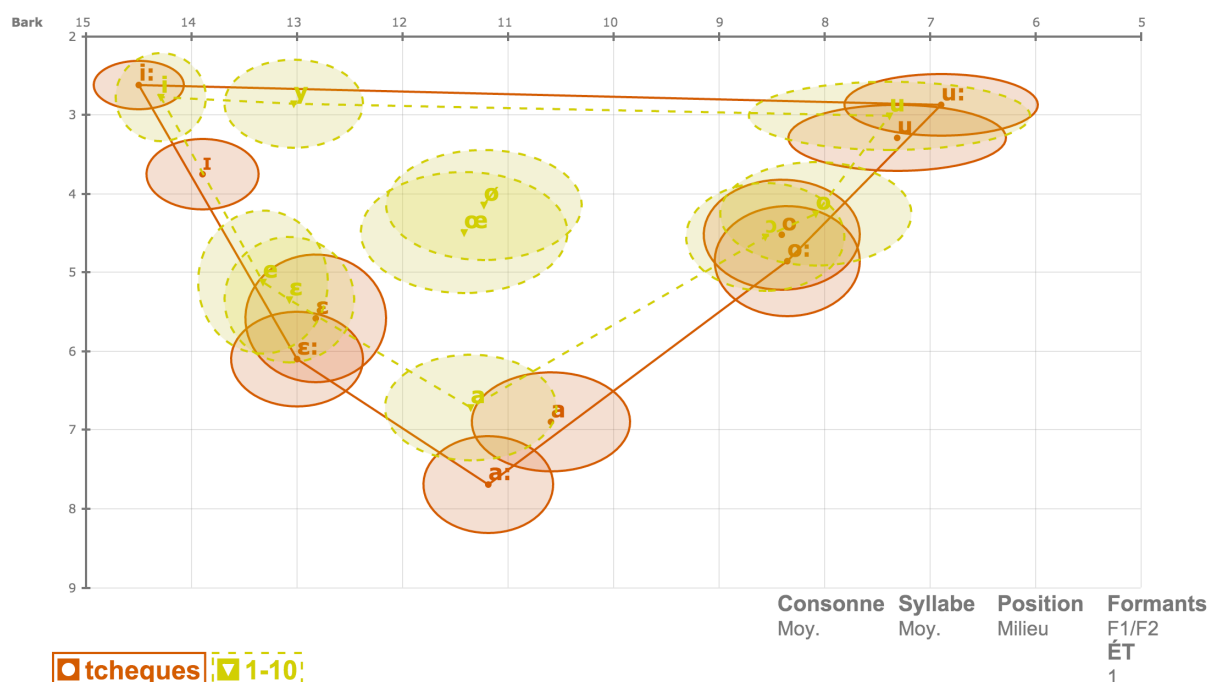


Figure 152 : Valeurs moyennes de F1/F2 (en Bark) des voyelles orales du français produites dans des logatomes CVCVCVC (C = p, t, k, R ; tous contextes confondus) par dix futures enseignantes (marquées par un triangle) et des voyelles du tchèque produites dans des logatomes CVCVCVC (C = p, t, k, h ; tous contextes confondus) par 20 Tchèques (marquées par un rond), 4 répétitions. Les ellipses de dispersion sont tracées à un écart-type

Ainsi les [i], [u], [a], [o, ɔ] et [e, ε] du français réalisés par les tchécophones sont, respectivement, proches par les formants de leurs [i:], [u/u:], [a], [o/o:] et [ε] tchèques. A l'inverse, les voyelles [y], [ø] et [œ] du français sont réalisées par des catégories phonétiques *nouvelles* (loin des sons maternels dans l'espace acoustique F1/F2).

L'identification par des Français natifs des voyelles isolées produites par les tchécophones montre que :

- les voyelles isolées [i, u, a, ε] sont produites avec authenticité ;
- [e] est largement confondu avec [ε] ;
- les voyelles des paires [o, ɔ] et [ø, œ] sont confondues entre elles.

Le résultat de l'identification par les futures enseignantes de FLE des voyelles produites par une Française est en accord avec leur performance en production, à l'exception de [ε] qui est largement confondu par les tchécophones avec [e].

L'ensemble des résultats révèle une difficulté évidente des futures enseignantes tchèques à maîtriser le timbre précis de chacune des six voyelles moyennes.

Puisque le but de notre travail était également de tester la fiabilité du SLM qui est le modèle théorique le plus adapté pour prédire le pattern de l'apprentissage chez des apprenants tardifs du niveau avancé, nous allons maintenant en discuter les limites. Quelques résultats vont dans la direction de l'impossibilité de l'acquisition authentique des voyelles *similaires* aux sons de la LM et au contraire de la maîtrise native des voyelles *nouvelles* (Flege and Hillenbrand, 1984). Cependant, certains résultats ne sont pas en accord avec le SLM. Il s'agit notamment des motifs d'apprentissage suivants :

- L'acquisition authentique de [i]. Cette voyelle du français est clairement *similaire* au [i:] long tchèque, sans être *identique* ni « hautement similaire ». Cependant, elle est produite et perçue authentiquement. Cette précision n'est pas prévue par SLM, certes, mais elle est en accord avec de nombreuses études montrant que les voyelles hautes sont généralement perçues sans difficulté, et cela même dans des tests d'assimilation interlangue (Strange *et al.*, 2005), et produites en LE fidèlement malgré le fait qu'elles ne sont pas *identiques* aux voyelles de la LM (Flege, 1997; Munro, 1993).
- La voyelle [ɛ], étant « hautement similaire » aux [ɛ/ɛ:] dans de nombreux contextes, est produite avec une grande précision, comme attendu. En revanche, les futures enseignantes de FLE ne possèdent pas de catégorie perceptive pour cette voyelle car elles la confondent largement avec /e/. Ce pattern d'apprentissage n'est pas prévu par Flege and Eefting (1987) pour qui les catégories perceptives pour un son étranger sont établies en premier lieu. Ce résultat confirme alors que la relation entre la perception et la production de sons de L2 est complexe, comme cela a été mis en évidence par exemple par Rochet (1995) ou Llisterri (1995, p. 98) selon qui : « *It does not seem possible to infer production abilities from perceptual ones and vice-versa.* »
- Les voyelles *similaires* ne sont pas toujours réalisées au travers de la catégorie maternelle phonétiquement la plus proche, comme le prédit SLM. Ainsi le [e] français n'est pas réalisé par les tchécoslovaques avec les caractéristiques acoustiques du [ɪ] bref tchèque qui est son correspondant acoustique et perceptif le plus proche. Il est plutôt réalisé et perçu par les tchécoslovaques au travers de la qualité vocalique [ɛ]. Ce phénomène peut être expliqué par l'influence négative de la graphie : le symbole [e] de même que la graphème <e> se prononcent en tchèque avec la qualité [ɛ]. Par conséquent, si les futures enseignantes sont exposées à des supports écrits, leur perception est troublée par la graphie et elles ne sont plus sensibles à la similarité du [e] avec le [ɪ] bref tchèque. Le même problème survient lors de l'apprentissage du [o] français qui est perceptivement « hautement similaire » au [u] tchèque. Sous l'influence de la graphie, les tchécoslovaques le produisent avec les caractéristiques acoustiques du [o] tchèque, donc avec un F1 plus élevé, ce qui conduit vers sa confusion par les natifs avec le /ɔ/. L'immense influence de la graphie a été étudiée par exemple par Escudero and Wanrooij (2010) et, auparavant, a déjà été commentée par Flege (1987a), (1987c) qui ne l'a néanmoins pas incorporée dans les hypothèses de base de SLM.
- Contre les attentes, une catégorie nouvelle est créée pour la voyelle [œ] qui est *similaire* parfois au niveau du pattern formantique et toujours perceptivement aux [ɛ/ɛ:] tchèques. L'inhibition de la création d'une nouvelle catégorie pour les sons analogues est expliquée dans SLM par le mécanisme cognitif de *classification par équivalence* (Flege, 1987d). En nous basant sur ce dernier, le [œ] du français devrait être perçu et réalisé par les tchécoslovaques par la qualité tchèque [ɛ]. L'on peut de nouveau expliquer ce pattern d'apprentissage inattendu par la graphie : du fait d'un symbole phonétique nouveau : [œ], correspondant aux graphèmes <oeu> ou <eu> qui n'existent pas ou ne sont jamais prononcés [ɛ] en tchèque, les tchécoslovaques échappent au mécanisme de la *classification par équivalence*. Dans ce cas, l'influence de la graphie est positive. Cependant [œ] n'est maîtrisé avec authenticité dans aucun contexte ; en revanche sa précision augmente dans les contextes où il est moins *similaire* au [ɛ] tchèque, ce qui est en accord avec SLM.
- Une catégorie nouvelle est créée pour la voyelle *nouvelle* [ø], comme le prédit SLM, mais cette voyelle n'est ni perçue ni produite avec authenticité par les futures enseignantes tchèques, ce qui est à priori contre les attentes. Il est important de rappeler que SLM prédit la possibilité

d'atteindre l'authenticité pour les sons *nouveaux* mais cela ne veut pas dire que l'authenticité est obligatoirement atteinte. En effet, la performance effective en LE chez des apprenants avancés s'explique par de nombreux facteurs, comme exposé ci-dessus. Puisque Flege a construit ses hypothèses à partir de l'analyse de la parole des apprenants avancés ayant souvent longtemps vécu dans le pays où l'on parle la L2, on peut supposer que la quantité et la qualité d'input de la LE que ces derniers ont obtenues sont en moyenne supérieures à celles de nos futures enseignantes de FLE. Nos sujets ont en effet principalement appris le français en milieu exolingue avec un enseignant non-natif. Ainsi le manque d'authenticité dans la maîtrise de sons *nouveaux* n'est pas dans l'absolu contre les hypothèses de SLM et doit être interprété en lien avec les facteurs agissant sur le niveau phonétique effectif.

Les résultats de notre étude révèlent que les prédictions concernant l'apprentissage phonétique d'une LE peuvent être établies dans un premier temps sur la base des symboles phonétiques et affinées par la suite par une analyse acoustico-perceptive. Au vu des résultats, les hypothèses de SLM devraient également prendre en compte une facilité générale dans l'acquisition des voyelles hautes et l'immense influence de la graphie. Ces deux points, s'ils ne sont pas explicités dans les postulats, peuvent en effet remettre en question le fondement du modèle SLM.

Une autre remarque peut être dégagée à partir de notre étude : l'analyse acoustique des productions des apprenants devrait être toujours accompagnée de tests perceptifs pour observer l'effet sur les auditeurs natifs. En effet, quoique la voyelle isolée [o] soit produite par les futures enseignantes avec des formants à moins d'un écart type de la « référence » (mais avec une distance F1-F2 supérieure à la « référence »), elle est perceptivement confondue par les Français natifs avec /ɔ/. De même, le [œ] isolé, produit à moins d'un écart type de la « référence », est confondu par les natifs avec /ø/. Il sera donc nécessaire, dans les études futures, de vérifier le timbre de toutes les voyelles produites par les futures enseignantes, y compris celui des voyelles insérées en contexte phonétique. En effet, sans vérification perceptive, les notions *authenticité* ou *maîtrise* ne devraient pas être utilisées.

Enfin, nos résultats rejoignent la position du SLM vis-à-vis de la période critique pour l'apprentissage : une telle période n'existe pas car l'apprentissage continue à un âge avancé. La création de nouvelles catégories en production et perception chez les Tchèques ayant appris le français après l'âge de 11 ans en témoignent. En revanche, l'effet de la LM et de la graphie sur la production et perception des voyelles de LE est immense et une fois que le système phonético-phonologique de la LM est formé et que l'apprentissage se fait avec des supports écrits, les sons sont interprétés en relation avec ces derniers.

La difficulté des futures enseignantes tchèques à percevoir et produire le contraste entre les voyelles moyennes e/ɛ, o/ɔ, ø/œ a été étudiée plus en détail dans une étude qualitative. Pour les dix futures enseignantes considérées, prises individuellement, nous avons testé la capacité à discriminer ces contrastes et analysé dix répliques de chacune des six voyelles qu'elles ont produites en isolation et en syllabe finale pVp, tVt, kVk et RVR. Les résultats de cette étude qualitative nous amènent vers les conclusions suivantes :

- Il est possible, mais rare, de surmonter l'influence de la graphie : la future enseignante T9 produit les voyelles [e] et [o] avec authenticité, comme en témoignent les résultats de l'analyse formantique et de leur identification par des auditeurs français natifs. Par ailleurs puisqu'il ne s'agit ni de voyelles *nouvelles* ni de voyelles (quasi)-*identiques* (ou « hautement similaire »), ce résultat semble en contradiction avec SLM. Or comme le précise Flege (1991), le SLM n'explique pas la variabilité dans la performance des apprenants. Ainsi, comme les neuf autres futures enseignantes ne maîtrisent pas ces voyelles, nos résultats ne remettent pas – vraiment – en question ce modèle.
- Certaines futures enseignantes tchèques possèdent deux catégories mentales bien distinctes pour les voyelles de chaque couplet ce qui se reflète dans leurs productions (ce résultat est en accord avec l'hypothèse H7 de SLM (Flege, 1995). Selon le contexte, les voyelles de chaque

paire s'opposent soit par l'indice acoustique principal, soit par l'indice acoustique secondaire de la « référence ».

- Au vu des résultats de l'identification par un Français natifs des voyelles synthétisées avec Klatt, la distance acoustique sur le plan de F1 et/ou F2 entre les voyelles de chaque paire produites par les tchécophones n'est probablement pas phonétiquement pertinente. Puisque l'oreille humaine est incapable de percevoir tous les ajustements phonétiques, il s'avère nécessaire de procéder à des analyses acoustiques fines si l'on veut évaluer objectivement le progrès dans la production des contrastes entre les voyelles moyennes. Par ailleurs, pour visualiser les résultats, nous avons créé le logiciel VisuVo qui génère en temps réel des graphiques à partir d'une grande base de données acoustiques.
- Lorsque les contrastes ne sont pas complètement acquis (c'est-à-dire lorsque les dix répliques de chaque catégorie ne diffèrent pas acoustiquement toutes l'une de l'autre), leur apprentissage peut néanmoins être en cours. Il se manifeste alors par des réalisations acoustiques vacillantes des sons représentant une catégorie. Ce résultat contredit Flege (1991) pour qui les sons étrangers *similaires* aux sons maternels sont peu variables car ils sont généralement produits avec la qualité acoustique des catégories maternelles bien établies et précises.
- Certains contextes phonétiques semblent faciliter la réalisation des contrastes entre les voyelles moyennes par les tchécophones. La notion de *contexte facilitant* telle que nous l'employons ici ne correspond pas à la conception classique des verbo-tonalistes qui placent les voyelles dans un contexte éclaircissant (par exemple [tyt]) ou assombrissant (par exemple [pyp]) selon la perception de l'apprenant (qui est influencée par sa LM). Le contexte facilitant tel que nous l'utilisons est défini par le résultat des tests de perception interlangue et c'est celui dans lequel les voyelles moyennes en contraste sont perceptivement assimilées à deux catégories tchèques différentes (ou l'un assimilé à une catégorie maternelle et l'autre à aucune catégorie). Il s'agit par exemple du contraste $\emptyset/\text{œ}$ en contexte dental où $[\emptyset]$ est assimilé aux $[u/u:]$ tchèques et $[\text{œ}]$ aux $[\text{ɛ}/\text{ɛ:}]$ tchèques. Dans ce cas précis de prédiction de la hiérarchie des difficultés dans la discrimination des contrastes entre les voyelles moyennes en fonction du contexte et pour définir les contextes facilitant l'apprentissage, le modèle PAM s'avère pertinent.
- Le lien entre la perception et production est loin d'être évident, comme exposé ci-dessus. Alors que la future enseignante T9 produit les voyelles isolées $[\text{e}]$ et $[\text{ɛ}]$ avec authenticité (comme en témoignent l'analyse formantique et l'identification par des Français natifs), elle les discrimine dans seulement deux suites sur 16. A l'inverse la future enseignante T10 discrimine ces voyelles dans 14 suites sur 16 mais échoue dans la production de la qualité $[\text{e}]$, qui est alors pratiquement toujours identifiée par les Français natifs en termes de $[\text{ɛ}]$.
- Enfin l'impossibilité d'atteindre la cible de $[\text{e}]$ en syllabe finale CVC et de $[\text{ɔ}]$ et $[\text{œ}]$ isolés par les tchécophones ne devrait pas être considérée comme de l'échec dans l'apprentissage phonétique des voyelles moyennes du français. En effet elle peut s'expliquer non seulement par le caractère *similaire* de ces voyelles du français avec des voyelles tchèques, comme le propose SLM, mais également par les règles orthoépiques. Il est possible que les futures enseignantes tchèques connaissent ces règles et les appliquent en produisant une qualité mi-ouverte en CVC et une qualité mi-fermée en isolation.

L'étude de la perception et production chez les Français natifs non-méridionaux a révélé que les contrastes entre les voyelles moyennes $\text{e}/\text{ɛ}$, $\text{o}/\text{ɔ}$, $\emptyset/\text{œ}$ sont marginaux car ils servent à opposer en français actuel un petit nombre d'unités sémantiques. Cette fragilisation va en effet dans le sens de la réduction phonologique la plus couramment employée dans les langues du monde, à savoir la neutralisation de contrastes d'aperture et notamment la suppression de voyelles moyennes dans les langues à trois degrés d'aperture, comme en russe, ou la suppression du contraste entre les voyelles mi-ouvertes et mi-fermées dans les langues à quatre degrés d'aperture, comme en italien (Flemming, 2006). Il est sans doute vrai que les contrastes avec un faible rendement fonctionnel sont en général moins importants à maîtriser que les contrastes très productifs car le nombre de contextes où l'inversion de la qualité peut diminuer la compréhensibilité et l'intelligibilité du message est restreint (Brown, 1991). En revanche, lorsqu'on veut 'faire français', il faut maîtriser en syllabe finale des mots

prosodiques les six timbres avec la même précision que les autres timbres de l'inventaire vocalique du français et les distribuer selon des règles bien précises.

Enfin, cette thèse a des implications didactiques sur l'apprentissage des voyelles de LE. Elle met en évidence que l'acquisition de la prononciation des voyelles isolées, dont la production est souvent illustrée dans les manuels de FLE par des schémas articulatoires, est insuffisante. Ce résultat appuie la position de Rvachew and Jamieson (1995) selon qui l'entraînement phonétique basé sur des voyelles isolées est peu efficace car une fois insérés dans des mots, les caractéristiques acoustiques des voyelles ne sont plus les mêmes. Et puisque « *the acquisition of coarticulation is not an automatic consequence of the acquisition of the phonological inventory.* » (Flege, 1997, p. 41), il faut prendre en compte les différents effets de coarticulation entre la LM et la LE et acquérir les règles allophoniques de la LE.

Les résultats de ce travail nous amènent en conclusion vers la formulation de conseils précis, adressés aux tchécophones souhaitant acquérir la prononciation native standard des voyelles orales du français actuel :

- **Faire attention aux différences de coarticulation entre le tchèque et le français.** Une fois les voyelles insérées dans la chaîne de la parole, le F2 des voyelles postérieures et du [a] français s'élève par rapport à la cible davantage que celui des voyelles correspondantes tchèques. L'élévation de F2 du [a] en contextes dental et palato-vélaire est si grande que les Tchèques monolingues l'assimilent aux voyelles tchèques de couleur « claire » [ɛ/ɛ:] (et non plus aux [a/a:] tchèques comme en isolation ou en contexte labial). De même, l'antériorisation du [ɔ] en [œ] en est un autre exemple (De Mareüil *et al.*, 2010).
- **Prendre conscience et connaissance de six qualités bien distinctes en syllabe finale d'un mot prosodique correspondant aux voyelles moyennes [o, ɔ, ø, œ, e, ɛ].**
- **Acquérir les six qualités précises des voyelles moyennes.** Les contextes facilitant l'acquisition des contrastes, définis par la perception interlangue, peuvent être utilisés pour rendre la différence acoustique entre les deux voyelles de chaque paire perceptivement plus saillante pour les oreilles tchèques.
- **Renforcer les oppositions entre les voyelles moyennes afin qu'elles soient perçues par les natifs.** Le contraste entre e/ɛ et ø/œ est essentiellement basé sur l'indice acoustique de F1, corrélé à l'aperture : [ɛ, œ] sont réalisés avec F1 plus élevé (corrélé à une plus grande aperture) que [e, ø]. Le F2 impacte ensuite la précision du timbre : [e] étant acoustiquement plus diffus se réalise avec F1 plus bas et F2 plus élevé que [ɛ] étant acoustiquement plus compact. Le contraste entre o/ɔ se base sur F1 et F2 (corrélé à l'antériorité-postériorité et labialité) : il est donc nécessaire de réaliser le [o] avec F1 et F2 plus bas et collés (corrélé à une plus grande fermeture et arrondissement des lèvres et postériorité de la langue) que ceux du [ɔ].
- **Veiller à ne pas enfreindre les règles orthoépiques dans la distribution des voyelles** selon le contexte consonantique (non étudiées dans notre thèse mais résumées par exemple par Léon (1976) et selon la structure syllabique. En syllabe finale CVC, réaliser toujours [ɛ] ; en syllabe finale CV, réalisez [ø] et [o].
- **Prendre connaissance des quelques cas particuliers de contextes où les contrastes sont toujours actifs.** Dans ces paires minimales, la qualité effective ne peut plus être distribuée selon la loi de position mais doit être réalisée d'après les règles particulières de prononciation indiquées dans le dictionnaire.
- **Enfin, dans les autres cas**, afin de ne pas alourdir l'apprentissage par des règles obsolètes, il est possible de suivre la tendance de la distribution complémentaire (voyelle fermée en syllabe ouverte et vice versa), notamment pour [e] en syllabe finale CV. Par ce point, nous rejoignons le modèle didactique de Wioland (2005) qui préconise la fermeture du timbre des voyelles moyennes en syllabe finale ouverte.

Bibliographie

- Abry, D. and Veldeman-Abry, J. (2007). *La phonétique : audition, prononciation, correction*. CLE, Paris.
- Adda-Decker, M., De Mareüil, P. B. and Lamel, L. (1999). Pronunciation variants in French : schwa & liaison. In: *ICPhS*, San Francisco
- Adu Manyah, K. and Sock, R. (2002). La quantité vocalique en twi. Considérations phonologiques et analyses acoustiques préliminaires. *XXIV Journées d'Etudes sur la Parole de l'Association Francophone de la Communication Parlée (JEP de l'AFCP)*, Nancy.
- Anderson-Hsieh, J., Johnson, R. and Koehler, K. (1992). The Relationship Between Native Speaker Judgments of Nonnative Pronunciation and Deviance in Segmentals, Prosody, and Syllable Structure. *Language Learning*, **42**, 529-555.
- Apostol, L., Perrier, P. and Bailly, G. (2004). A model of acoustic interspeaker variability based on the concept of formant-cavity affiliation. *Journal of the Acoustical Society of America*, **115**, 337-351.
- Arai, T. (2001). The replication of Chiba and Kajiyama's mechanical models of the human vocal cavity. *Journal of Phonetic Society of Japan*, **5**, 31-38.
- Arai, T. (2004). History of Chiba and Kajiyama and their influence in modern speech science. *From Sound to Sense: 50+ Years of Discoveries in Speech Communication*, Boston.
- Asher, J. J. and Garcia, R. (1969). The optimal age to learn a foreign language. *The Modern Language Journal*, **311**, 334-341.
- Ašić, T. (2008). *Espace, temps, prépositions*. Librairie Droz S. A., Genève.
- Asla Ostby, K. (2005). Le système vocalique du français parisien. Une étude dans le cadre du projet "Phonologie du français contemporain". In: *Présentation donnée à l'IRIT (Institut de Recherche en Informatique de Toulouse)*, Toulouse.
- Atal, B. S., Chang, J. J., Mathews, M. V. and Tukey, J. W. (1978). Inversion of articulatory-to-acoustic transformation in the vocal-tract by a computer-sorting technique. *Journal of the Acoustical Society of America*, **63**, 1535-1555.
- Azoulay-Vicente, A. (1990). Pour un enseignement intégré de la phonétique. *Revue de Phonétique Appliquée*, **94**, 1-15.
- Badin, B., Perrier, B., Boë, L.-J. and Abry, C. (1990). Vocalic Nomograms : Acoustic and articulatory considerations upon formant convergences. *Journal of the Acoustical Society of America*, **78**, 1290-1300.
- Baken, R. J. and Orlikoff, R. F. (2000). *Clinical Measurement of Speech and Voice*. Singular, United States of America.
- Bell-Berti, F., Raphael, L. J., Pisoni, D. B. and Sawusch, J. R. (1979). Some Relationships between Speech Production and Perception. *Phonetica*, **36**, 373-383
- Bent, T. and Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, **114**.
- Besse, H. (1985). *Méthodes et pratiques des manuels de langue*. Didier-Crédif, Paris.
- Best, C. T. (1995). A Direct Realistic View of Cross-Language Speech Perception. In: *Speech Perception and Linguistic Experience : Theoretical and Methodological Issues in Cross-Language Speech Research*, W. Strange, Ed, Timonium, MD : York Press Inc.

- Best, C. T., Hallé, P., Bohn, O. S. and Faber, A. (2003). Cross-language perception of nonnative vowels: Phonological and phonetic effects of listeners' native languages. *ICPhS*, Barcelone, Espagne, pp. 649-663.
- Best, C. T., McRoberts, G. W. and Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, **109**.
- Best, C. T., McRoberts, G. W., Lafleur, R. and Silver-Isenstadt, J. (1995). Divergent Developmental Patterns for Infants' Perception of Two Non-Native Consonant Contrasts. *Infant Behaviour and development*, **18**, 339-350.
- Best, C. T., McRoberts, G. W. and Sithole, N. M. (1988). Examination of Perceptual Reorganization for Nonnative Speech Contrasts: Zulu Click Discrimination by English-Speaking Adults and Infants. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 345-360.
- Best, C. T. and Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In: *Second language speech learning: The role of language experience in speech perception and production*, M. J. Munro and O. S. Bohn, Eds, pp. 13-34.
- Billières, M. (1987). Le suivi des élèves en phonétique corrective. *Revue de Phonétique Appliquée*, **82-83-84**, 15-28.
- Billières, M. and Borrell, A. (1990). Quelques problèmes soulevés par les différentes variétés d'accents dans les méthodes de français langue étrangère. *Revue de Phonétique Appliquée*, **94**.
- Billières, M. and Spanghero-Gaillard, N. (2005). La didactique cognitive des langues: regards croisés de disciplines pour comprendre le "comment". *Revue Parole*, **34-35-36**.
- Birdsong, D. (2003). Authenticité de prononciation en français L2 chez des apprenants tardifs anglophones : analyses segmentales et globales. In: *Acquisition et interaction en langue étrangère*.
- Blake, E. S. (1972). A Self-Correction Checklist for French Vowels. *The French Review*, **46**, 59-66.
- Blanchet, P. (1993). Voyelles moyennes et accent tonique en français de Provence. *La Linguistique*, **29**.
- Boersma, P. and Weenink, D. (1992-2011). Praat: doing phonetics by computer, <http://www.praat.org>, Ed.
- Bohn, O. S. and Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, **11**, 303-328.
- Bohn, O. S. and Flege, J. E. (1997). Perception and production of a new vowel category by second-language learners. In: *Second-language speech: Structure and process*, A. James and J. Leather, Eds, Walter de Gruyter, Berlin, pp. 53-74.
- Bongaerts, T., van Summeren, C., Planken, B. and Schils, E. (1997). Age and Ultimate Attainment in the Pronunciation of a Foreign Language. *Studies in Second Language Acquisition*, **19**.
- Borovičková, B. and Maláč, V. (1967). *The spectral analysis of Czech sound combinations*. Praha.
- Borrell, A. (1993). Relation entre les aspects articulatoires et les aspects acoustiques en phonétique. Quels outils utiliser en didactique des langues ? *Revue de Phonétique Appliquée*, **1993**.
- Borrell, A. (1996). Systématisation des erreurs de production et donc de perception chez les apprenants étrangers ? *Revue de Phonétique Appliquée*, **118-119**.

- Borrell, A. and Billières, M. (1989). L'évolution de la norme phonétique en français contemporain. *La Linguistique*, **25**, 45-62.
- Bothorel, A., Simon, P., Wioland, F. and Zerling, J.-P. (1986). *Cinéradiographie des voyelles et consonnes du français*. Travaux de l'Institut de Phonétique de Strasbourg, Strasbourg.
- Brainard, M. S. and Doupe, A. J. (2002). What songbirds teach us about learning. *Nature*, **417**, 351-358.
- Brosnahan, L. F. and Malmberg, B. (1970). *Introduction to Phonetics*. W. Heffer & Sons Ltd, Cambridge.
- Brown, A. (1991). Functional load and the teaching of pronunciation. In: *Teaching English Pronunciation : A book of readings*, A. Brown, Ed, Routledge, London, pp. 211-224.
- Burnham, D. K. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics*, **7**, 207-240.
- Calliope. (1989). La parole et son traitement automatique. In: *Collection Technique et Scientifique des Télécommunications*, J. P. Tubach, Ed, Masson, Paris.
- Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B. and Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America*, **54**.
- Carlson, R., Granstrom, B. and Fant, G. (1970). Some studies concerning perception of isolated vowels. *Speech Transmission Laboratory: Q. Prof. Stat. Rep.*, **2-3**, 19-35.
- Catford, J. C. and Pisoni, D. B. (1970). Auditory vs Articulatory Training in Exotic Sounds. *The Modern Language Journal*, **54**, 477-481.
- Cechova, M., Dokulil, M., Hlavsa, Z., Hrbacek, J. and Hruskova, Z. (2000). *Čeština - řeč a jazyk*. ISV, Praha.
- Chistovich, L. A. and Lubilnskaya, V. V. (1979). The 'Center of gravity' effect in vowel spectra and critical distance between the formants : Psychoacoustical study of the perception of vowel-like stimuli. *Hearing Research*, **1**, 185-195.
- Chladkova, K. and Podlisky, V. J. (2011). Native dialect matters : Perceptual Assimilation of Dutch vowels by Czech listeners. *Journal of the Acoustical Society of America*, **130**.
- Chlumsky, J. (1928). *Ceska kvantita, melodie a prizvuk*. Ceska akademie, Praha.
- Clarke, C. M. and Garrett, M. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, **116**.
- Conseil de l'Europe, C. E. (2001). Un cadre européen commun de référence pour les langues: apprendre, enseigner, évaluer, Didier.
- Constantinescu-Stefanel, R. (2006). Les méthodologies non-conventionnelles. *Dialogos*, **14**, 74-79.
- CREDIF. (1972). *De vive voix*. Didier, Paris.
- Crothers, J. (1978). Typology and universals of vowel systems. In: *Universals of Human Language*, J. Greenberg, C. A. Ferguson and E. A. Moravcsik, Eds, Vol. 2, Stanford University Press, Stanford, California.
- Cunningham-Andersson, U. and Engstrand, O. (1989). Perceived Strength and Identity of Foreign Accent in Swedish. *Phonetica*, **46**, 138-154.
- Cureau, J. (1973). Le Suvag Lingua de classe. Note d'expérimentation. *Revue de Phonétique Appliquée*, **27-28**, 35-42.
- Dankovicova, J. (1997a). Czech. *Journal of the International Phonetic Association*, **27**, 77-80.

- Dankovicova, J. (1997b). The domain of articulation rate variation in Czech. *Journal of Phonetics*, **25**, 387-312.
- Dankovicova, J. and Dellwo, V. (2007). Czech Speech Rhythm and the Rhythm Class Hypothesis. In: *ICPhS XVI*, Saarbrücken.
- Darcy, I., Dekydtspotter, L., Sprouse, R. A., Glover, J., Kaden, C., McGuire, M. and Scott, J. H. G. (2012). Direct mapping of acoustics to phonology: On the lexical encoding of front rounded vowels in L1 English- L2 French acquisition. *Second Language Research*, **28**, 5-40.
- Dart, S. N. (1998). Comparing French and English coronal consonant articulation. *Journal of Phonetics*, **26**, 71-94.
- De Mareüil, P. B., Adda-Decker, M. and Woehrling, C. (2010). Antériorisation/aperture des voyelles / O/~o/ en français du Nord et du Sud. *28es Journées d'Étude sur la Parole*, Mons.
- Delattre, P. (1939). Accent de mot et accent de groupe. *The French Review*, **13**, 141-146.
- Delattre, P. (1944). Vers la méthode phonétique intégrale pour les débutants. *The French Review*, **18**, 109-115.
- Delattre, P. (1945). Prononciation Graphique et Prononciation Phonétique: II. Les Voyelles. *The French Review*, **18**, 285-296.
- Delattre, P. (1947). *An Introduction to Franch Speech Habits*. Holt, Rinehart and Winston, inc., United States.
- Delattre, P. (1948). Phonetics in Beginning Language Study. *The Modern Language Journal*, **32**, 373-377.
- Delattre, P. (1951). The Physiological Interpretation of Sound Spectograms. *PMLA*, **66**, 864-875.
- Delattre, P. (1953). Les modes phonétiques du français. *The French Review*, **27**, 59-63.
- Delattre, P. (1960). Sommes-nous préparés pour l'enseignement de la langue? *The French Review*, **33**, 483-490.
- Delattre, P. (1963). Le jeu des transitions de formants et la perception des consonnes. *Proc. IVth Internal. Congr. phonetic Sciences*, La Haye, pp. 407-418.
- Delattre, P. (1964). Comparing the vocalic features of English, German, Spanich and French. *International Review of Applied Linguistics in Language Teaching*, **2**.
- Delattre, P. (1968). La Radiographie des voyelles françaises et sa correlation acoustique. *The French Review*, **42**, 48 - 65.
- Delattre, P. (1969). An Acoustic and Articulatory Study of Vowel Reduction in Four Languages *Internation Review of Applied Linguistics in Language Teaching*, **7**.
- Delattre, P., Liberman, A. M. and Cooper, F. S. (1951). Voyelles synthétiques à deux formantes et voyelles cardinales. *Le Maître Phonétique*, **96**, 30-36.
- Delattre, P., Liberman, A. M., Cooper, F. S. and Gertsman, L. J. (1952). An experimental study of the acoustic determinants of vowel color : observation on the one- and two-formant vowels sythetised from spectrographic patterns. *Word*, **3**, 195-210.
- Delvaux, V., Metens, T. and Soquet, A. (2002). Propriétés acoustiques et articulatoires des voyelles nasales du français. *Actes des 24èmes Journées d'Etude sur la Parole*, Nancy, pp. 357-360.
- Derwing, M. and Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, **42**, 476-490.
- Derwing, T. M., Munro, M. J. and Wiebe, G. (1998). Evidence in Favor of a Broad Framework for Pronunciation Instruction. *Language Learning*, **48**, 393-409.

- Detey, S., Durand, J., Laks, B. and Lyche, C. (2010). Les variétés du français parlé dans l'espace francophone. Ressources pour l'enseignement, Ophrys, Paris.
- Detey, S. and Racine, I. (2012). Les apprenants de français face aux normes de prononciation : Quelle(s) entrée(s) pour quelle(s) sortie(s) ? *Revue française de linguistique appliquée*, **17**, 81-96.
- Dickerson, L. B. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly*, **9**, 401-407.
- Dobre, D. (2006a). La méthode audio-orale. *Dialogos*, **14**.
- Dobre, D. (2006b). La méthode directe. *Dialogos*, **14**.
- Dohalska, M., Dubeda, T., Bartosova, H. and Mejvaldova, J. (2000). Spectral Properties of Czech Vowels in Spontaneous Speech (Preliminary Analysis). In: *10th Czech-German Workshop*, Praha.
- Dohalska, M. and Schulzova, O. (2003). *Fonetika francouzštiny*. Karolinum, Praha.
- Dohalska-Zichova, M. (1980). Perception et analyse spectrographique des voyelles tchèques et françaises. *Acta Universitatis Carolinae, Phonetica Pragensia*, **6**, 97-116.
- Dommergues, J.-Y. (2008). Rôle des contraintes phonologiques sur la reconnaissance des mots : l'exemple de deux dialectes du français. In: *Actes de la 2eme conférence de l'ISPhS*, U. C. d. Prague, Ed, Karolinum, Praha, pp. 63-73.
- Dubeda, T. (2002). Unité accentuelle en français et en tchèque. Caractérisation structurelle et acoustique. In: *Institut de phonétique*, Université Charles, Prague.
- Dubeda, T. (2004). K izosylabičnosti a izochronnosti v češtině. *ISPhS*.
- Dubeda, T. (2005). *Jazyky a jejich zvuky : univerzálie a typologie ve fonetice a fonologii*. Karolinum, Praha.
- Dubeda, T. and Januska, J. (2006). K hypoartikulaci v pražské češtině. In: *Sborník příspěvků z 5. mezinárodní konference Setkání mladých lingvistů*, P. Porizka and V. P. Polach, Eds, Univerzita Palackého v Olomouci, Olomouc, pp. 75-81.
- Dubois, J. and Lagane, R. (1995). *Grammaire*. Larousse, Milan.
- Duchet, J.-L. (1981). *La phonologie*. PUF, Paris.
- Duez, D. (2001). Manifestation phonétique de la réduction et de l'assimilation contextuelle des segments de la parole conversationnelle. *Parole*, **17-18-19**, 89-111.
- Duez, D. and Dubeda, T. (2006). Les faits de réduction et d'assimilation dans l'enseignement du français : pour une "phonétique situationnelle". In: *La classe de langue – théories, méthodes et pratiques*, M. Faraco, Ed, Publications de l'Université de Provence, pp. 231-248.
- Durand, J. (A paraître). L'alphabet phonétique international. In: *Le monde des écritures*, C. Herrenschmidt, R. Mugnaioni, M. J. Savelli and C. Touratier, Eds, Gallimard, Paris.
- Durand, P. (1985). *Variabilité acoustique et invariance en français : consonnes occlusives et voyelles*. Paris.
- Dvorak, V. (2010). Voicing assimilation in Czech. In: *Rutgers Working Papers in Linguistics*, P. Staroverov, Ed, Vol. 3, Rutgers University, NJ : LGSA, pp. 115-144.
- Eckman, F. R. (1977). Markedness and The Contrastive Analysis Hypothesis. *Language Learning*, **27**, 315-330.
- Eckman, F. R. (1985). Some Theoretical and Pedagogical Implications of the Markedness Differential Hypothesis. *Studies in Second Language Acquisition*, **7**, 289-307.

- Eckman, F. R., Elreyes, A. and Iverson, G. K. (2001). Allophonic Splits in L2 Phonology : The Question of Learnability. *International Journal of English Studies*, **1**.
- Eckman, F. R., Elreyes, A. and Iverson, G. K. (2003). Some principles of second language phonology. *Second Language Research*, **19**, 169-208.
- Eckman, F. R. and Iverson, G. K. (2013). The Role of Native Language Phonology in the Production of L2 Contrasts. *Studies in Second Language Acquisition*, **35**, 67-92.
- Eddy, F. D. (1974). Pierre Delattre, Teacher of French. *The French Review*, **47**, 513-517.
- Els, v. and Bot, d. (1987). The Role of Intonation in Foreign Language. *The Modern Language Journal*, **71**, 147-155.
- Enoch, P. (1967). Etude expérimentale sur la perception de phonèmes français par des auditeurs israéliens. *Revue de Phonétique Appliquée*, **4**, 13-20.
- Escudero, P. (2001). The role of the input in the development of L1 and L2 sound contrasts: language-specific cue weighting for vowels. In: *Proceedings of the 25th Annual Boston University Conference on Language Development*, A. H.-J. Do, L. Dominguez and A. Johansen, Eds, Cascadilla.
- Escudero, P. and Boersma, P. (2001). The subset problem in L2 perceptual development: multiple-category assimilation by Dutch learners of Spanish. *Proceedings*, **24**.
- Escudero, P. and Polka, L. (2003). A Cross-language Study of Vowel Categorization and Vowel Acoustics : Canadian English versus Canadian French. In: *ICPhS*, Barcelona.
- Escudero, P. and Wanrooij, K. (2010). The Effect of L1 Orthography on Non-native Vowel Perception. *Language and Speech*, **53**, 343-365.
- Esling, J. H. (1987). Methodology for Voice Setting Awareness in Language Classes. *Revue de Phonétique Appliquée*, **85**.
- Fagyal, Z., Hassa, S. and Ngom, F. (2002a). L'opposition [e]-[E] en syllabes ouvertes de fin de mot en français parisien : étude acoustique préliminaire. In: *XXIVème Journées d'Etude sur la parole*, Nancy.
- Fagyal, Z., Nguyen, N. and De Mareüil, P. B. (2002b). From dilation to coarticulation : is there vowel harmony in French ? *Studies in the Linguistic Sciences*, **32**.
- Fant, G. (1960). *The Acoustic Theory of Speech Production*. Mouton, The Hague.
- Fant, G. (1973). *Speech sounds and features*. MIT Press, Cambridge.
- Fant, G. (1983). Feature analysis of Swedish vowels - A revisit. *KTH, Speech Transmission Laboratory - Quarterly Progress ans Status Report*, **2-3**, 1-19.
- Fenclova, M. (2003). *Fonetika francouzštiny jako lingvodidaktický problém*. Univerzita Karlova v Praze - Pedagogická fakulta, Praha.
- Fernandez, J. G. (2012). L'enseignement de la prononciation : rapport entre théorie et pratique. *Revue française de linguistique appliquée*, **17**, 67-80.
- Féry, C. (2011). Markedness, Faithfulness, Vowel Quality and Syllable Structure in French. *Linguistics in Potsdam*, **15**.
- Flanagan, J. L. (1955). A Difference Limen for Vowel Formant Frequency. *Journal of the Acoustical Society of America*, **27**.
- Flanagan, J. L. (1972). *Speech Analysis, Synthesis and Perception*. Springer-Verlag, Berlin.
- Flaux, N. (1993). *La grammaire*. Presses Universitaires de France, Paris.

- Flege, J. E. (1980). Phonetic Approximation In Second Language Acquisition. *Language Learning*, **30**.
- Flege, J. E. (1981). The Phonological Basis of Foreign Accent : A Hypothesis. *TESOL Quarterly*, **15**, 443-455.
- Flege, J. E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, **76**.
- Flege, J. E. (1987a). A Critical Period for Learning to Pronounce Foreign Languages ? *Applied Linguistics*, **8**, 162-177.
- Flege, J. E. (1987b). The Instrumental Study of L2 Speech Production: Some Methodological Considerations. *Language Learning*, **37**.
- Flege, J. E. (1987c). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, **15**, 67-83.
- Flege, J. E. (1987d). The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, **15**, 47-65.
- Flege, J. E. (1988a). Factors affecting degree of perceived accent in English sentences. *Journal of the Acoustical Society of America*, **84**.
- Flege, J. E. (1988b). The Production and Perception of Foreign Language Speech Sounds. In: *Human Communication and Its Disorders, A Review*, H. Winitz, Ed, Ablex, Norwood, pp. 224-401.
- Flege, J. E. (1989). Chinese subjects' perception of the word-final English /t/-/d/ contrast: Performance before and after training. *Journal of the Acoustical Society of America*, **86**.
- Flege, J. E. (1991). Perception and production: the relevance of phonetic input to L2 phonological training. In: *Crosscurrents in Second Language Acquisition and Linguistic Theories*, K. Elenius and P. Branderud, Eds, John Benjamins, Amsterdam.
- Flege, J. E. (1995). Second Language Speech Learning. Theory, Findings, and Problems. In: *Speech Perception and Linguistic Experience : Theoretical and Methodological Issues in Cross-Language Speech Research*, W. Strange, Ed, Timonium, MD : York Press Inc.
- Flege, J. E. (1997). English vowel production by Dutch talkers: More evidence for the "similar" vs "new" distinction. In: *Second-language speech, Structure and process*, A. James and J. Leather, Eds, Mouton de Gruyter, Berlin, pp. 11-52.
- Flege, J. E. (2003). A method for assesing the perception of vowels in a second language. In: *Issues in Clinical Linguistics*, E. Fava and A. Mioni, Eds, UniPress, Padova, pp. 19-44.
- Flege, J. E. and Eefting, W. (1987). Cross-language switching in stop consonant perception and production by dutch speakers of English. *Speech Communication*, **6**, 185-202.
- Flege, J. E. and Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America*, **76**, 708-721.
- Flemming, E. (2006). *A phonetically-based model of phonological vowel reduction*. Ms., MIT.
- Flemming, E. and Johnson, S. (2007). Rosa's roses: reduced vowels in American English. *Journal of the International Phonetic Association*, **37**, 83-96.
- Fokes, J. and Bond, Z. S. (1989). The Vowels of Stressed and Unstressed Syllables in Nonnative English. *Language Learning*, **39**.
- Fonagy, I. (1980). L'accent en français : accent probabilitaire : dynamique d'un changement prosodique. In: *L'accent n français contemporain*, I. Fonagy and P. Léon, Eds.
- Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, **29**, 109-135.

- Fougeron, C., Gendrot, C. and Bürki, A. (2007). On the phonetic identity of French schwa compared to /ø/ and /œ/. In: *Journées d'Etudes Linguistiques*, Nantes, France.
- Fourakis, M. (1991). Tempo, stress, and vowel reduction in American English. *Journal of the Acoustical Society of America*, **90**.
- Fourakis, M. and Iverson, G. K. (1985). On the Acquisition of Second Language Timing Patterns. *Language Learning*, **35**.
- Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A. and Hallé, P. (2008). Cross language phonetic influences on the speech of French-English bilinguals. *Journal of Phonetics*, **36**, 649-663.
- Fraisse, P. (1974). *Psychologie du rythme*. PUF, Paris.
- Freland-Ricard, M. (1996). Organisation temporelle et rythmique chez les apprenants étrangers. Etude multilingue. *Revue de Phonétique Appliquée*, **118-119**, 61-89.
- Galazzi-Matasci, E. (2011). La correction phonétique aujourd'hui: permanences, ruptures, résurgences. Questionnement et perspectives en FLE. In: *Journées IPFC*, Paris.
- Galazzi-Matasci, E. and Pedoya, E. (1983). Et la pédagogie de la prononciation. *Le français dans le monde*, **180**, 39-44.
- Garde, P. (1968). *L'accent*. P.U.F., Paris.
- Gendrot, C. and Adda-Decker, M. (2005). Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German. *Interspeech*, pp. 2453-2456.
- Gendrot, C. and Adda-Decker, M. (2010). Influence du contexte consonantique et de la durée des voyelles sur la centralisation des voyelles orales en français. In: *La Coarticulation, des Indices aux Représentations*, M. Embarki and C. Dodane, Eds, Harmattan, Paris, pp. 159-171.
- Gendrot, C., Adda-Decker, M. and Vaissière, J. (2008). Les voyelles /i/ et /y/ du français : focalisation et variations formantiques. *JEP*, Avignon, pp. 205-208.
- Gendrot, C. and Gerdes, K. (2009). Prosodic hierarchy and spectral realization of vowels in French. In: *Actes d'IDP 2009*, H.-Y. Yoo and E. Delais-Roussarie, Eds, Paris.
- Georgeton, L., Paillereau, N., Landron, S., Gao, J. and Kamiyama, T. (2012). Analyse formantique des voyelles orales du français en contexte isolé : à la recherche d'une référence pour les apprenants de FLE. In: *XXIXièmes Journées d'Etude sur la Parole*, Grenoble, pp. 145-152.
- Goldman, J.-P. (2011). EasyAlign : an automatic phonetic alignment tool under Praat. *InterSpeech*, Firenze, Italy.
- Gottfried, T. L. (1984). Effects of consonant context on the perception of French Vowels. *Journal of Phonetics*, **12**, 91-114.
- Gottfried, T. L. and Beddor, P. S. (1988). Perception of Temporal and Spectral Information in French Vowels. *Language and Speech*, **31**, 57-74.
- Grammont, M. (1933). *Traité pratique de prononciation française*. Delagrave, Paris.
- Grepl, M., Furst, T. and Pesak, J. (2007). The F1-F2 Vowel Chart for Czech Whispered Vowels A, E, I, O, U. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub.*, **151**, 353-356.
- Grosjean, F. (2013). Retaining an Accent. Why some people retain an accent in second language. In: *Psychology today*.
- Grosjean, F., Carrard, S., Godio, C., Grosjean, L. and Dommergues, J.-Y. (2007). Long and short vowels in Swiss French : Their production and perception. *Journal of French Language Studies*, **17**, 1-19.

- Guberina, P. (1970). Phonetic Rhythms in the Verbo-Tonal System. *Revue de Phonétique Appliquée*, **16**, 3-13.
- Guberina, P. (1973). Les appareils Suvag et Suvag Lingua. *Revue de Phonétique Appliquée*, **27-28**, 7-15.
- Guberina, P., Gospodnetic, Y., Pozojevic, M., Skaric, J. and Vuletic, B. (1965). Correction de la prononciation des élèves qui apprennent le français. *Revue de Phonétique Appliquée*, **1**, 81-94.
- Guimbretière, E. (1994). *Phonétique et enseignement de l'oral*. Didier/Hatier, Paris.
- Guion, S. G., Flege, J. E., Akahane-Yamada, R. and Pruitt, J. S. (2000). An investigation of current models of second language speech perception : the case of Japanese adult's perception of English consonants. *Journal of the Acoustical Society of America*, **107**.
- Hála, B. (1941). *Akustická podstata samohlásek*. Česká akademie věd a umění, Praha.
- Hallé, P., Segui, J., Dominguez, L. and Cuertos, F. (a paraître). *special* is *especial* but *stuto* is not *astuto*: Perception of prothetic /e/ in speech and print by speakers of Spanish. In: *Psicolingüística en Español. Homenaje a Juan Seguí*, V. Jaichenco and Y. Sevilla, Eds, Secretaría de Publicaciones, Facultad de Filosofía y Letras, Universidad de Buenos Aires, Buenos Aires.
- Hallé, P., Segui, J., Frauenfelder, U. and Meunier, C. (1998). The processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 592-608.
- Harmegnies, B., Delvaux, V., Huet, K. and Piccaluga, M. (2005). Oralité et cognition: pour une approche raisonnée de la pédagogie du traitement de la matière phonique. *Revue Parole*, **34-35-36**.
- Hillenbrand, J., Clarke, C. M. and Nearey, T. M. (2001). Effects of consonant environment on vowel formant patterns. *Journal of the Acoustical Society of America*, **109**, 748-763.
- Hillenbrand, J. and Gayvert, R. T. (1993). Vowel Classification Based on Fundamental Frequency and Formant Frequencies. *Journal of Speech and Hearing Research*, **36**, 694-700.
- Hirsch, F., Sock, R., Roy, J.-P. and Canault, M. (2004). Perception anticipatoire catégorielle des voyelles arrondies du français. *XXV Journées d'Etudes sur la Parole de l'Association Francophone de la Communication Parlée (JEP de l'AFCP)*, Fès.
- Hoensch, J. K. (1995). *Histoire de la Bohême*. Payot, Paris.
- Honda, K., Kitamura, T., Takemoto, H., Mokhtari, P. and Adachi, S. (2008). Resonance characteristics of the hypopharyngeal cavities. *Journal of the Acoustical Society of America*, **123**.
- Howie, S. M. (2001). Formant transitions of Russian palatalized and nonpalatalized syllables. *IULC Working Papers Online 1*.
- Huckvale, M. (2009). VTDemo, University College London.
- International Phonetic Association. (2010). The Principles of the International Phonetic Association (1949). *Journal of the International Phonetic Association*, **40**, 299-358.
- Intravaia, P. (1976). Pour une méthodologie de l'investigation en phonétique corrective. *Revue de Phonétique Appliquée*, **38**.
- Intravaia, P. (1993). Pour une étiologie approfondie de l'erreur phonétique. *Revue de Phonétique Appliquée*, **108-109**, 239-263.
- Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. and Siebert, C. (2001). A perceptual interference account of acquisition difficulties for non-native phonemes. *Speech, Hearing and Language: work in progress*, **13**.

- Johnson, K. (1997). *Acoustic and Auditory Phonetics*, Blackwell.
- Kamiyama, T. and Vaissière, J. (2009). Perception and production of French close and close-mid rounded vowels by Japanese-speaking learners. *Acquisition et interaction en langue étrangère*, **2**.
- Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *Journal of the Acoustical Society of America*, **67**, 971-975.
- Komarek, M. (1982). *Nástin fonologického vývoje českého jazyka*. Státní pedagogické nakladatelství, Praha.
- Krashen, S. (1975). The critical period for language acquisition and its possible bases. In: *Developmental Psycholinguistics and Communication Disorders*, D. Aaronson and R. Rieber, Eds, New York Academy of Sciences, New York, pp. 211-224.
- Kucera, H. and George, K. M. (1968). *A comparative quantitative phonology of russian, czech and german*. American Elsevier Publishing Company, New York.
- Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, **97**.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M. and Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society*, **363**, 979-1000.
- Kuhl, P. K. and Iverson, P. (1995). Linguistic Experience and the "Perceptual Magnet Effect". In: *Speech Perception and Linguistic Experience : Theoretical and Methodological Issues in Cross-Language Speech Research*, W. Strange, Ed, Timonium, MD : York Press Inc., pp. 121-150.
- Kundera, M. (1987). *La vie est ailleurs*. Gallimard, Paris.
- Lado, R. (1957). *Linguistics across cultures : applied linguistics for language teachers*. University of Michigan Press.
- Lambert-Drache, M. (1997). *Sur le bout de la langue: introduction au phonétisme du français*, Canadian Scholars' Press.
- Landercy, A. (1976). Méthodologie SGAV et système verbo-tonal: synthèse de tendances nouvelles. *Revue de Phonétique Appliquée*, **39-40**.
- Landercy, A. and Renard, R. (1976). Observations sur les fondements acoustiques et perceptifs de certains procédés du système verbo-tonal. *Revue de Phonétique Appliquée*, **19-40**, 175-194.
- Landick, M. (1995). The Mid-Vowels in Figures: Hard Facts. *The French Review*, **69**, 88-102.
- Landron, S., Paillereau, N., Nawafleh, A., Exare, C., Ando, H. and Gao, J. (2010). Le corpus PhoDiFLE : un corpus commun de français langue étrangère pour une étude phonétique des productions de locuteurs de langues maternelles plurielles. *Cahiers de Praxématique*, **54-55**, 73-86.
- Leather, J. (1983). Second-language pronunciation learning and teaching. *Language Teaching*, **16**, 198-219.
- Leather, J. and James, A. (1991). The Acquisition of Second Language Speech. *Studies in Second Language Acquisition*, **13**, 305-341.
- Lebel, J. G. (1969). Apport des exercices systématiques de discrimination auditive dans la correction phonétique. *Revue de Phonétique Appliquée*, **10**, 3-9.
- Lefebvre, A. (1988). Les voyelles moyennes dans le français de la radio et de la télévision. *La Linguistique*, **24**.

- Lenneberg, E. (1967). *Biological Foundations of Language*. Wiley, New York.
- Léon, P. R. (1976). *Prononciation du français standard*. Didier, Paris.
- Léon, P. R. (1993). *Phonétisme et prononciation du français*.
- Léon, P. R. and Léon, M. (2007). *La prononciation du français*. Armand Colin, Paris.
- Léon, P. R. and Tennant, J. (1990). Indices de perceptibilité et de différenciabilité des timbres vocaliques : la variabilité [e] - [ɛ] en français. *Revue québécoise de linguistique*, **19**, 9-22.
- Levy, E. S. (2009). Language experience and consonantal context effects on perceptual assimilation of French vowels by American-English learners of French. *Journal of the Acoustical Society of America*, **125**.
- Levy, E. S. and Strange, W. (2008). Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics*, **36**, 141-157.
- Liljencrants, J. and Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, **48**, 839-862.
- Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, **35**, 1773-1781.
- Lindblom, B. (1978). Final lengthening in speech and music. In: *Nordic Prosody*, E. Garding, Ed, Lund, pp. 85-101.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H and H theory. In: *Speech production and speech modeling*, W. J. Hardcastle and A. Marchal, Eds, Kluwer, Dordrecht, pp. 403-439.
- Llisterri, J. (1995). Relationships between speech production and speech perception in a second language. In: *ICPhS*, Stockholm, Sweden.
- Ludvikova, M. and Kraus, J. (1966). Kvantitativni vlastnosti soustavy ceskych fonemu. *Slovo a slovesnost*, **XXVII**, 334-344.
- MacKain, K. S. (1982). Assessing the role of experience on infants' speech discrimination. *Journal of Child Language*, **9**, 527-542.
- MacKain, K. S., Best, C. T. and Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, **2**, 369-390.
- Maddieson, I. and Wright, R. (1996). Small vowel systems and phonetic variability – evidence from Amis. *Austronesian Formal Linguistics Association*, **3**.
- Maeda, S. (1982). A digital simulation method of the vocal-tract system. *Speech Communication*, **1**, 199-229.
- Maeda, S. (1989a). Compensatory Articulation during Speech. In: *Speech Production and Modelling*, W. J. Hardcastle and A. Marchal, Eds, Academic Publishers, Kluwer, pp. 131-149.
- Maeda, S. (1989b). Compensatory Articulation during Speech: Evidence from the Analysis and Synthesis of Vocal-Tract Shapes using an Articulatory Model. In: *Speech Production and Modelling*, W. J. Hardcastle and A. Marchal, Eds, Kluwer Academic Publishers.
- Maeda, S. and Carré, R. (1996). Modèles de production. In: *Fondements et perspectives en traitement automatique de la parole*, H. Méloni, Ed, Aupelf Uref.
- Magnen, C., Billières, M. and Gaillard, P. (2005). Surdit  phonologique et cat gorisation. Perception des voyelles fran aises par les hispanophones. *Revue Parole*, **33**, 9-34.
- Major, R. C. (1987). Phonological similarity, markedness, and rate of L2 acquisition. *Studies in Second Language Acquisition*, **9**, 63-82.

- Manuel, S. Y. (1990). The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Haskins Status Report on Speech Research*.
- Manuel, S. Y. and Krakow, R. A. (1984). Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskins Status Report on Speech Research*, **77-78**.
- Martin, P. (2008). Phonétique acoustique. Introduction à l'analyse acoustique de la parole, Armand Colin, Paris.
- Martinet, A. (1945). *La Prononciation du français contemporain , témoignages recueillis en 1941 dans un camp d'officiers prisonniers*. Paris, E. Droz
- Martinet, A. and Walter, H. (1973). *Dictionnaire de la prononciation française dans son usage réel*. France-Expansion, Paris.
- Mayerthaler, W. (1988). *Morphological Naturalness*. Karoma publishers, Ann Arbor, USA.
- Mazlova, V. (1946). Jak se projevuje zvukova stranka cestiny v hlaskovych statistikach ? *Nase rec*, **30**, 101-111.
- McAllister, R. (1998). Second Language Perception and the Concept of Foreign Accent. In: *ETRW on Speech Technology in Language Learning*, Marholmen, Sweden.
- McAllister, R. (2000). Perceptual foreign accent and its relevance for simultaneous interpreting. In: *Language Processing and Simultaneous Interpreting*, K. Hytlenstam and B. Englund-Dimitrova, Eds, John Benjamin, Amsterdam.
- McCarthy, B. (1991). L'accentuation en classe de français langue étrangère. *Revue de Phonétique Appliquée*, **98**.
- Ménard, L., Chrétien, J., Lachapelle, R. and Marieau, I. (2010). Corrélat des voyelles chez les Sourds. *Spectrum*, **2**.
- Métral, J. P. (1967). Phonétique et phonologie. *Revue de Phonétique Appliquée*, **4**, 21-47.
- Meunier, C. (2001). Le traitement de la variabilité dans la parole. Aspects théoriques et méthodologiques. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence*, **20**, 69-90.
- Meunier, C. (2007). Phonétique acoustique. In: *Les dysarthries*, P. Auzou, Ed, Masson, pp. 164-173.
- Meunier, C., Espesser, R. and Frenck-Mestre, C. (2006). Aspects phonologique et dynamique de la distinctivité au sein des systèmes vocaliques : une étude inter-langue. In: *Actes des 26èmes Journées d'études sur la Parole*, Dinard, pp. 333-336.
- Miller, J. L., Grosjean, F., Mondini, M. and Dommergues, J.-Y. (2011). Dialect effects in speech perception : The role of vowel duration in Parisian French and Swiss French. *Language and Speech*, **54**, 467-485.
- Miller, L. J. (1981). Effects of speaking rate on segmental distinctions. In: *Perspectives on the Study of Speech*, P. D. Eimas and J. L. Miller, Eds, Erlbaum, NJ.
- Mooshammer, C. and Geng, C. (2008). Acoustic and articulatory manifestations of vowel reduction in German. *Journal of the International Phonetic Association*, **38**, 117-136.
- Morin, Y. C. (1983). Pour une histoire des voyelles longues en français : quelques problèmes. In: *The Linguistic Association of the Atlantic Provinces*, University of Moncton.
- Moulton, W. G. (1962). Toward a Classification of Pronunciation Errors. *The Modern Language Journal*, **46**, 101-109.
- Munro, M. J. (1993). Production of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech*, **36**.

- Munro, M. J. and Derwing, T. M. (1995a). Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning*, **45**, 73-97.
- Munro, M. J. and Derwing, T. M. (1995b). Processing Time, Accent, and Comprehensibility in the Perception of Native and Foreign-Accented Speech. *Language and Speech*, **38**.
- Murillo, J. (1982). Le seuil de phonologisation. *Revue de Phonétique Appliquée*, **64**.
- Nemser, W. (1993). The vocalic repertoires of two australian learners of English. In: *Current Issues in European Second Language Acquisition Research*, B. Kettemann and W. Vieden, Eds, gnv, Tübingen.
- Öhman, S. E. G. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *Journal of the Acoustical Society of America*, **40**, 979-988.
- Ouni, S. and Laprie, Y. (2001). Exploring the Null Space of the Acoustic-to-Articulatory Inversion Using a Hypercube Codebook. *Eurospeech2001*, Aalborg, Denmark.
- Oyama, S. (1976). A Sensitive Period for the Acquisition of a Nonnative Phonological System. *Journal of Psycholinguistic Research*, **5**.
- Palková, Z. (1997). *Fonetika a fonologie češtiny*. Karolinum, Praha.
- Palmen, M. J., Bongaerts, T. and Schils, E. (1997). L'authenticité de la prononciation dans l'acquisition d'une langue étrangère au-delà de la période critique : des apprenants néerlandais parvenus à un niveau très avancé en français. *Acquisition et interaction en langue étrangère*, **9**.
- Peterson, G. E. and Barney, H. L. (1952). Control Methods Used in a Study of the Vowels. *Journal of the Acoustical Society of America*, **24**, 175-184.
- Piske, T., MacKay, I. R. A. and Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, **29**, 191-215.
- Podlipsky, V. J. (2009). The Interaction between Word-Boundary Perception and Vowel Length in Native and Non-Native Speakers of Czech. *Acta Univ Palackianae Olomucensis, Fac Philos, Anglicka III, Linguistica*, 89-100.
- Podlipsky, V. J., Skarnitzl, R. and Volin, J. (2009). High Front Vowels in Czech: a Contrast in Quantity or Quality? In: *Interspeech*, Brighton, UK, pp. 132-135.
- Pognan, P. (1999). Histoire de l'écriture et de l'orthographe tchèques. *Histoire - Epistémologie - Langages*, **21**, 27-62.
- Polivanov, E. (1931). The perception of non-native language sounds. *Travaux du Cercle Linguistique de Prague*, **4**, 79-96.
- Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic and acoustic contributions. *Journal of the Acoustical Society of America*, **89**.
- Polka, L. and Werker, J. F. (1994). Developmental Changes in Perception of Nonnative Vowel Contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, **20**, 421-435.
- Pullum, G. K. and Ladusaw, W. A. (1996). *Phonetic Symbol Guide*. The University of Chicago, Chicago.
- Puren, C. (1988). *HISTOIRE DES MÉTHODOLOGIES DE L'ENSEIGNEMENT DES LANGUES*. Nathan-CLE International, Paris.
- Puren, C. (1990). Méthodes d'enseignement, méthodes d'apprentissage et activités méta-méthodologiques en classe de langue. *Langues Modernes, revue de l'APLV*, **1**.
- Puren, C. (2001). Méthodes et constructions méthodologiques dans l'enseignement et

- l'apprentissage des langues *Les Langues modernes*, **1**, 68-70.
- Puren, C. (2006). De l'approche communicative à la perspective actionnelle. *Le français dans le monde*, **347**, 37-40.
- Pustka, E. (2011). L'accent méridional : représentations, attitudes et perceptions toulousaines et parisiennes. *Varia*, **69**.
- Racine, I., Detey, S. and Kawaguchi, Y. (2012). Les voyelles /y-u/ dans IPFC : évaluation perceptive de productions natives, hispanophones et japonophones. *JEP-TALN-RECITAL*, Grenoble, pp. 385-392.
- Rakova, Z. (2011). Francophonie de la population tchèque. In: *La francophonie en Europe centrale et pour l'Europe centrale*, M. Fenclova and D. Kolarikova, Eds, Zapadoceska univerzita, Plzen
- Ricci, J. F. A. (1936). Fantaisie sur le pur Accent. *The Modern Language Journal*, **20**, 363-365.
- Rigault, A. (1970). L'accent dans deux langues à accent fixe: le français et le tchèque. *Studia Phonetica*, **3**.
- Riordan, C. J. (1977). Control of vocal-tract length in speech. *Journal of the Acoustical Society of America*, **62**, 4998-1002.
- Rochet, B. L. (1995). Perception and Production of Second-Language Speech Sounds by Adults. In: *Speech Perception and Linguistic Experience : Theoretical and Methodological Issues in Cross-Language Speech Research*, W. Strange, Ed, Timonium, MD : York Press Inc., pp. 379-407.
- Rossi, M. (1980). Le français, langue sans accent ? *Studia Phonetica*, **15**, 85-101.
- Rousset, I. (2004). Structures syllabiques et lexicales des langues du monde Données, typologies, tendances universelles et contraintes substantielles. In: *Linguistique*, Université Stendhal - Grenoble III.
- Rvachew, S. and Jamieson, D. G. (1995). Learning New Speech Contrasts. Evidence from Adults Learning a Second language and Children with Speech Disorders. In: *Speech Perception and Linguistic Experience : Theoretical and Methodological Issues in Cross-Language Speech Research*, W. Strange, Ed, Timonium, MD : York Press Inc.
- Saunders, G. E. (2009). Why phonetics matters: food for thought for aspiring teachers of French. In: *Phonetics Teaching & Learning Conference*, UCL, London.
- Scovel, T. (1978). The Effect of Affect on Foreign Language Learning : A Review of the Anxiety Research. *Language Learning*, **28**, 129-141.
- Shelly, S. L. (1998). Teaching the Sound System(s): The Case of Mid Vowels. *The French Review*, **71**, 598-606.
- Simackova, S., Podlipsky, V. J. and Chladkova, K. (2012). Czech spoken in Bohemia and Moravia. *Journal of the International Phonetic Association*, **42**, 225-232.
- Skarnitzl, R. (2014). Asymmetry in the Czech alveolar stops. An EPG study. *Acta Universitatis Carolinae Philologica*, **1**.
- Skarnitzl, R. and Volin, J. (2012). Referenční hodnoty vokálních formantů pro mladé dospělé mluvčí standardní češtiny. *Akustické listy*, **18**, 7-11.
- Southwood, M. H. and Flege, J. E. (1999). Scaling foreign accent: direct magnitude estimation versus interval scaling. *Clinical Linguistics and Phonetics*, **13**, 335-349.
- Stalhammar, U., Karlsson, I. and Fant, G. (1973). Contextual affects on vowel nuclei. *STL-QPSR*, **14**, 1-18.

- Stevens, K. N. (1989). On the Quantal Nature of Speech. *Journal of Phonetics*, **17**.
- Stevens, K. N. (1998). *Acoustic Phonetics*. MIT Press, Cambridge.
- Stevens, K. N. and House, A. S. (1963). Perturbation of Vowel Articulations by Consonantal Context : An Acoustical Study. *Journal of Speech and Hearing Research*, **6**, 111-128.
- Stevens, K. N., Libermann, A. M., Studdert-Kennedy, M. and Ohman, S. E. G. (1969). Crosslanguage Study of Vowel Perception. *Language and Speech*, **12**.
- Stoean, C.-S. (2006). La méthode traditionnelle. *Dialogos*, **14**.
- Straka, G. (1964). L'évolution phonétique du latin au français sous l'effet de l'énergie et de la faiblesse articulatoire. *T. L. L., Centre de Philologie Romane, Strasbourg*, **2**, 17-28.
- Strange, W. (1989). Evolving theories of vowel perception. *Journal of the Acoustical Society of America*, **85**.
- Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A. and Nishi, K. (2001). Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *Journal of the Acoustical Society of America*, **109**.
- Strange, W. and Bohn, O. S. (1998). Dynamic specification of coarticulated German vowels: Perceptual and acoustical studies. *Journal of the Acoustical Society of America*, **104**.
- Strange, W., Bohn, O. S., Nishi, K. and Trent, S. A. (2005). Contextual variation in the acoustic and perceptual similarity of North German and American English vowels. *Journal of the Acoustical Society of America*, **118**.
- Strange, W., Jenkins, J. J. and Johnson, T. L. (1983). Dynamic specification of coarticulated vowels. *Journal of the Acoustical Society of America*, **74**, 523-351.
- Syrdal, A. K. and Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English Vowels. *Journal of the Acoustical Society of America*, **79**.
- Tahal, K. (2010). *A grammar of Czech as a Foreign Language*. Factum CZ.
- Tahta, S., Wood, M. and Leowenthal, K. (1981). Foreign Accents : Factors Relating to Transfer of Accent from the first Language to a Second Language. *Language and Speech*, **24**.
- Taylor, L. L., Catford, J. C., Guiora, A. Z. and Lane, H. (1971). Psychological Variables and Ability to Pronounce a Second Language. *Language and Speech*, **14**.
- Trofimovich, P., Baker, W. and Mack, M. (2001). Contexte- and Experience-based effects on the learning of vowels in a second language. *Studies in the Linguistic Sciences*, **31**.
- Trubetzkoy, N. S. (1957). *Principes de phonologie*. Klincksieck, Paris.
- Trubetzkoy, N. (1939). Grundzüge der Phonologie. *Travaux du Cercle Linguistique de Prague*, **7**, 1-271.
- Urbain, W. (1975). Suvaglingua et pose spontanée de la voix. *Revue de Phonétique Appliquée*, **33-34**, 81-106.
- Vaissière, J. (1996). From Latin to Modern French: On Diachronic Changes and Synchronic Variations. *AIPUK. Arbetisberitche, Institut fur Phonetik und digitale Sprechverarbeitung, Universitat Kiehl*, **31**, 61-74.
- Vaissière, J. (2001). Changements de sons et changements prosodiques : du latin au français. *Revue Parole*, **17/18/19**, 53-88.
- Vaissière, J. (2006). *La phonétique*. Presses universitaires de France, Paris, pp. 1 vol. (125).

- Vaissière, J. (2007). Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of the contrast between the sounds in a language. In: *Experimental Approaches to Phonology*, P. S. Beddor, M. J. Solé and J. J. Ohala, Eds, Oxford University Press, Oxford, pp. 54-72.
- Vaissière, J. (2009). Articulatory modeling and the definition of acoustic-perceptual targets for reference vowels. *The Chinese Phonetics Journal*, **2**, 22-33.
- Vaissière, J. (2010). Le français, langue à frontières par excellence. In: *Frontières: du linguistique au sémiotique*, N. Andrieux-Reix, Ed, Lambert-Lucas.
- Vaissière, J. (2011). On the Acoustic and Perceptual Characterization of Reference Vowels in a Cross-Language Perspective. In: *ICPhS XVII*, Hong Kong.
- Valdman, A. (2000). Comment gérer la variation dans l'enseignement du français langue étrangère aux Etats-Unis. *The French Review*, **73**, 648-666.
- van Bergem, D. R. (1993). Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech Communication*, **12**, 1-23.
- van Bergem, D. R. (1994). Reflections on Aspects of Vowel Reduction. *Proceedings of the Institute of Phonetic Sciences*, **18**.
- van Wijngaarden, S. J., Steeneken, H. J. M. and Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native talkers. *Journal of the Acoustical Society of America*, **112**.
- Volin, J. and Studenovsky, D. (2007). Normalization of Czech Vowels from continuous read texts. In: *ICPhS*, Saarbrücken.
- Wachs, S. (2011). Tendances actuelles en enseignement de la prononciation du français, langue étrangère (FLE). *Revista de Linguas Modernas*, **14**, 183-196.
- Walter, H. (1976). *La dynamique des phonèmes dans le lexique français contemporain*. France-Expansion.
- Walter, H. (1988). *Le français dans tous les sens*. Robert Laffont, Paris.
- Weinreich, U. (1953). *Languages in contact, findings and problems*. Linguistic Circle of New York, New York.
- Wen, B. J. and Wioland, F. (1982). Is French really syllable-timed ? *Journal of Phonetics*, **10**, 193-216.
- Wioland, F. (1991). *Prononcer les mots du français : des sons et des rythmes*. Hachette, Paris.
- Wioland, F. (2001). Que faire de la graphie "e" dans le cadre de la didactique de la prononciation ? . *Le français dans le monde*, **318**.
- Wioland, F. (2005). *La vie sociale des sons du français*. L'Harmattan, Paris.
- Wode, H. (1994). Nature, Nurture, and Age in Language Acquisition. *Studies in Second Language Acquisition*, **16**, 325-345.
- Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*, **24**, 245-261.
- Zwicker, E. and Fastl, H. (1990). *Psychoacoustics: Facts and Models*. Springer-Verlag, Berlin.

Liste des tableaux

Tableau 1 : Similarité spectrale des voyelles antérieures arrondies de l'allemand du nord (NG) et des voyelles postérieures de l'anglais américain (AE) en fonction du contexte, selon Strange <i>et al.</i> (2005). % indique le pourcentage de voyelles antérieures de l'allemand qui correspondent aux voyelles postérieures de l'anglais dans un contexte donné	37
Tableau 2 : Les motifs d'assimilation perceptive des voyelles de l'anglais américains (AE) en termes de catégories de l'allemand du nord (NG), tout contexte confondu (colonne 3-4), en contextes b-p (colonne 5-6), d-t (colonne 7-8) et g-k (colonne 9-10), selon Strange <i>et al.</i> (2005)	38
Tableau 3 : Les motifs d'assimilation des voyelles de l'anglais américain (AE) aux voyelles japonaises sans prendre en compte leur durée et la palatalisation, tout contexte consonantique confondu, selon Strange <i>et al.</i> (2001) ..	39
Tableau 4 : L'assimilation perceptive des voyelles du français parisien (PF) en termes de catégories anglaises (AE) en contextes labial (à gauche) et coronal (à droite), selon Levy (2009).....	39
Tableau 5 : Etudes antérieures sur les différences de stratégies de coarticulation entre les langues.....	41
Tableau 6 : Assimilation perceptive des voyelles de l'anglais en termes de voyelles du coréen, selon Trofimovich <i>et al.</i> (2001)	43
Tableau 7 : Le « difference limen » de F1 et F2, selon Flanagan (1955)	46
Tableau 8 : Paires minimales attestant du caractère phonémique des voyelles du tchèque	52
Tableau 9 : Paires minimales attestant du caractère phonémique des voyelles du français.....	53
Tableau 10 : Tableau comparatif des voyelles du tchèque et du français.....	54
Tableau 11 : Valeurs moyennes des formants F1, F2, F3, F4 pour les dix voyelles orales du français [i, e, ε, a, u, o, ɔ, y, ø, œ], selon Georgetown <i>et al.</i> (2012)(GD), calculées à partir de 1600 voyelles isolées (10 voyelles*4 répétitions*40 locutrices) et Gendrot et Adda-Decker (2005) (G&A), calculées à partir de 2 heures de parole radiophonique (15 femmes, tout contexte consonantique). Les écarts types sont entre parenthèses (avec les valeurs extrêmes grisées), N/D : non défini.	55
Tableau 12 : Les zones de formants F1 et F2 pour les voyelles brèves, selon Grepl <i>et al.</i> , et les moyennes formantiques (avec les écarts types entre parenthèses) pour les voyelles brèves et longues du tchèque de 27 hommes, selon Skarnitzl et Volin (S&V). N/D signifie non défini.....	59
Tableau 13 : Les valeurs formantiques des voyelles tchèques, en Hertz, en ERB et la différence en pourcentage entre les valeurs formantiques des voyelles brèves et longues au niveau de chaque paire. Les écarts types sont entre parenthèses. Selon Skarnitzl and Volin (2012)	60
Tableau 14 : Système consonantique du tchèque ; les phonèmes inexistants dans le système français sont en gras ..	61
Tableau 15 : Paires minimales attestant du caractère phonémique des consonnes du tchèque	61
Tableau 16 : Système consonantique du français ; les phonèmes qui ne font pas partie du système consonantique tchèque sont en gras	62
Tableau 17 : Paires minimales attestant du caractère phonémique des consonnes du français.....	62
Tableau 18 : Paires minimales attestant du caractère phonémique des semi-consonnes du français.....	62
Tableau 19 : Tableau comparatif des consonnes du tchèque et du français.....	63
Tableau 20 : Etudes antérieures sur l'apprentissage phonétique du français	77
Tableau 21 : Modèles récents de l'apprentissage phonétique	85
Tableau 22 : Méthodologies conventionnelles d'enseignement des langues étrangères	94
Tableau 23 : Informations concernant les locutrices françaises natives indiquant leur âge (avec la moyenne et l'écart type entre parenthèses), la région où elles ont grandi et le nombre d'années passées en Ile de France (avec la moyenne et l'écart type entre parenthèses).....	106
Tableau 24 : Informations concernant les futures enseignantes de FLE de LM tchèque indiquant l'âge au début de l'apprentissage du français, le nombre d'années d'étude du français (avec les moyennes et les écarts types entre parenthèses) et le nombre de mois passés en France.....	107

Tableau 25 : Informations concernant les locutrices tchèques natives indiquant leur âge (avec la moyenne et l'écart type entre parenthèses), la région où elles ont grandi et le nombre d'années passées en Bohême (avec la moyenne et l'écart type entre parenthèses).....	108
Tableau 26 : Critères acoustiques utilisés dans notre thèse pour définir les voyelles <i>identiques</i> , « hautement similaires », <i>similaires</i> et <i>nouvelles</i> de même que les voyelles authentiques/ non authentiques, non maîtrisées.....	119
Tableau 27 : Corpus, locutrices pour les mesures des moyennes formantiques des voyelles isolées du français et du tchèque.....	130
Tableau 28 : Valeurs moyennes (en Hz) de F1 et F2 (toutes les voyelles), F3 (voyelles antérieures) et F4 (voyelles [i, ɪ, i:]) des voyelles du français et du tchèque (3 valeurs*10 locutrices*4 répétitions). Les écarts types se trouvent entre parenthèses. Les valeurs extrêmes sont en gras. FR = français, TCH = tchèque	131
Tableau 29 : Distances acoustiques F2-F1 (de toutes les voyelles), F3-F2 (voyelles antérieures et [a]) et F4-F3 (voyelles [i, ɪ, i:]) des voyelles du français (FR) et du tchèque (TCH) en Hertz. Les rapprochements formantiques sont grisés et le rapprochement maximal est en gras. Les valeurs moyennes sont calculées à partir de productions de 10 locutrices*4 répétitions pour le français et 20 locutrices*4 répétitions pour le tchèque.....	135
Tableau 30 : Aperture 0 : Valeurs formantiques moyennes (en Hz) des voyelles françaises fermées (calculées à partir de 3 valeurs*10 Françaises *4 répétitions). Les écarts types (exprimés en Hz et en pourcentage) se trouvent entre parenthèses	137
Tableau 31 : Aperture 1 : Valeurs formantiques moyennes (en Hz) des voyelles françaises mi-fermées (calculées à partir de 3 valeurs*10 Françaises *4 répétitions). Les écarts types (exprimés en Hz et en pourcentage) se trouvent entre parenthèses	138
Tableau 32 : Aperture 2 : Valeurs formantiques moyennes (en Hz) des voyelles françaises mi-ouvertes (calculées à partir de 3 valeurs*10 Françaises *4 répétitions). Les écarts types (exprimés en Hz et en pourcentage) se trouvent entre parenthèses	139
Tableau 33 : Aperture 0 : Valeurs formantiques moyennes des voyelles tchèques fermées [i:, ɪ, u:, u] (calculées à partir de 3 valeurs*20 Tchèques *4 répétitions). Les écarts types (exprimés en Hz et en pourcentage) se trouvent entre parenthèses	140
Tableau 34 : Aperture 1 : Valeurs formantiques moyennes (en Hz) des voyelles moyennes du tchèque [ɛ:, ɛ, o:, o] (calculées à partir de 3 valeurs*20 Tchèques *4 répétitions). Les écarts types (exprimés en Hz et en pourcentage) se trouvent entre parenthèses.....	141
Tableau 35 : Les valeurs formantiques des voyelles tchèques, en Hz (1 ^e colonne), en ERB (2 ^e colonne) et la différence en pourcentage (3 ^e colonne) entre les valeurs formantiques des voyelles brèves et longues au niveau de chaque paire. Les écarts types sont entre parenthèses	148
Tableau 36 : Similarité acoustique entre les voyelles du français (FR) et du tchèque (TCH) prononcées en isolation, exprimée en termes d'écarts types (ET). <= 1 ET => Similarité +++ (gris foncé), 1 – 1,5 ET => Similarité ++ (gris moyen), 1,5 – 2 ET => Similarité + (gris clair), 2 et + ET => Similarité 0.....	151
Tableau 37 : Les mouvements des formants F1, F2, (F3) durant la palatalisation de [u, ʊ], labialisation de [ʌ, ɪ] et vélarisation de [i, y] par rapport à la cible. Deux flèches indiquent une variation formantique plus importante que celle des autres formants.....	155
Tableau 38 : Corpus, locutrices et mesures des moyennes formantiques des voyelles du français prononcées en contextes p, t, k, R.....	156
Tableau 39 : Distance acoustique (en Hertz) entre les valeurs moyennes formantiques des voyelles du français prononcées en contextes [p, t, k, ʁ] (calculées à partir de productions de 10 locutrices*4 répétitions*3 syllabes) et les valeurs des formants des mêmes voyelles prononcées en isolation (calculées à partir de productions de 10 locutrices*4 répétitions*3 valeurs). Les mouvements formantiques importants (supérieurs à 50 Hz pour F1 et à 95 Hz pour F2, F3 et F4) sont grisés	156
Tableau 40 : Résultat de l'ANOVA à 2 facteurs qui indique l'effet du contexte sur le formant F1 des dix voyelles orales du français (10 locutrices*4 répétitions*3 syllabes).....	160
Tableau 41 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Contexte et Voyelle*Contexte sur le F1 des voyelles orales du français (10 locutrices*4 répétitions*3 syllabes)	160
Tableau 42 : Résultat de l'ANOVA à 2 facteurs qui indique l'effet du contexte sur le formant F2 des dix voyelles orales du français (10 locutrices*4 répétitions*3 syllabes).....	162

Tableau 43 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Contexte et Voyelle*Contexte sur le F2 des voyelles du français (10 locutrices*4 répétitions*3 syllabes).....	163
Tableau 44 : Résultat de l'ANOVA à 2 facteurs qui mesure l'effet du contexte sur le formant F3 des dix voyelles orales du français (10 locutrices*4 répétitions*3 syllabes)	171
Tableau 45 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Contexte et Voyelle*Contexte sur le F3 des voyelles du français (10 locutrices*4 répétitions*3 syllabes).....	171
Tableau 46 : Distances acoustiques F2-F1 (de toutes les voyelles), F3-F2 (voyelles antérieures) et F4-F3 (voyelle [i]) en Hertz des voyelles en isolation (0) et en contextes p, t, k, R. Les moyennes formantiques calculées à partir de productions de 10 locutrices*4 répétitions (valeur prise à la moitié de la durée vocalique). Le contexte favorisant la focalisation est en gras, le contexte fortement défavorisant est grisé	174
Tableau 47 : Durée moyenne (en ms) des voyelles du français produites en isolation et en contextes p, t, k, R (4 répétitions*3 positions*10 locutrices). L'écart type est entre parenthèses	175
Tableau 48 : Les mouvements de F2 et F3 des voyelles antérieures étirées/ arrondies, voyelles postérieures et [a] en contextes p, t, k, R par rapport à la cible. ↑ signifie une augmentation de la valeur et ↓ indique sa baisse. La case grisée indique un résultat inattendu selon la modélisation dans VTDemo. Le nombre de flèches indique l'ampleur des mouvements formantiques (1 flèche par tranche de 100 Hz).....	176
Tableau 49 : Les mouvements de F2 et F3 des voyelles antérieures étirées et arrondies en contextes [p, t, k, ʁ] par rapport à la cible. ↑ signifie une augmentation de la valeur et ↓ indique sa baisse. Le nombre de flèches indique l'ampleur des mouvements formantiques (1 flèche par tranche de 100 Hz).....	176
Tableau 50 : Contextes consonantiques qui conservent les valeurs formantiques cibles de voyelles du français subséquentes.....	177
Tableau 51 : Corpus, locutrices et mesures des moyennes formantiques des voyelles du tchèque prononcées en contextes p, t, k, h	178
Tableau 52 : Distance acoustique (en Hertz) entre les valeurs moyennes formantiques centrales des voyelles tchèques prononcées en contextes [p, t, k, h], calculées à partir de productions de 20 locutrices*4 répétitions*3 syllabes et les valeurs formantiques moyennes des mêmes voyelles prononcées en isolation, calculées à partir de productions de 20 locutrices* 4 répétitions*3 valeurs. Les mouvements formantiques importants sont grisés	179
Tableau 53 : Résultat de l'ANOVA à 2 facteurs qui mesure l'effet du contexte sur le formant F1 des dix voyelles monophthongues du tchèque (20 locutrices*4 répétitions*3 syllabes).....	181
Tableau 54 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Contexte et Voyelle*Contexte sur le F1 des voyelles du tchèque (20 locutrices*4 répétitions*3 syllabes)	181
Tableau 55 : Résultat de l'ANOVA à 2 facteurs qui mesure l'effet du contexte sur le formant F2 des dix voyelles monophthongues du tchèque (20 locutrices*4 répétitions*3 syllabes).....	182
Tableau 56 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Contexte et Voyelle*Contexte sur le F2 des voyelles du tchèque (20 locutrices*4 répétitions*3 syllabes).....	182
Tableau 57 : Résultat de l'ANOVA à 2 facteurs qui mesure l'effet du contexte sur le formant F3 des dix voyelles monophthongues du tchèque (20 locutrices*4 répétitions*3 syllabes).....	186
Tableau 58 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Contexte et Voyelle*Contexte sur le F3 des voyelles du tchèque.....	187
Tableau 59 : Distances acoustiques F2-F1 (de toutes les voyelles), F3-F2 (voyelles antérieures) et F4-F3 (voyelles [i, i:]) en Hertz des voyelles du tchèque selon le contexte phonétique. Les moyennes formantiques sont calculées à partir de productions de 20 locutrices*4 répétitions (valeur prise à la moitié de la durée vocalique). Le contexte favorisant la focalisation est en gras, le contexte fortement défavorisant est grisé.....	188
Tableau 60 : Durée moyenne (en ms) des voyelles tchèques produites en isolation et en contextes p, t, k, h (4 répétitions*3 positions*20 locutrices). L'écart type est entre parenthèse	189
Tableau 61 : Les mouvements de F2 et F3 en fonction du groupe vocalique (voyelles antérieures/ postérieures, longues/brèves et [a]) et de la consonne subséquente (p, t, k, h). ↑ signifie une augmentation de la valeur et ↓ indique sa baisse.....	190
Tableau 62 : Contextes consonantiques qui conservent les valeurs formantiques cibles de voyelles subséquentes du tchèque.....	192

Tableau 63 : Similarité des voyelles du français (colonne à gauche) et du tchèque en contextes [p, t, k] et nul (0), précisé entre parenthèses. – [= 1 ET = Similarité +++, 1 – 1,5 ET = Similarité ++, 1,5 – 2 ET = Similarité +, 2 et + ET = Similarité 0	197
Tableau 64 : Corpus, locutrices et mesures des moyennes formantiques des voyelles du français et du tchèque prononcées dans 3 positions du logatomes	201
Tableau 65 : Résultat de l'ANOVA à un facteur mesurant l'effet de la syllabe sur la durée vocalique.....	202
Tableau 66 : PLSD de Fisher pour le facteur « Syllabe » avec niveau de significativité de 5 %, FIN = syllabe finale, MID = syllabe médiane, BEG = syllabe initiale	202
Tableau 67 : Valeurs moyennes de F1 et F2 (toutes les voyelles), F1 à F3 (voyelles antérieures) des voyelles du français, prononcées en syllabes initiale, médiane et finale dans des logatomes CVCVCVC où C = [p, t, k, ʁ] (valeur prise à la moitié de la durée* 4 répétitions* 4 contextes confondus* 10 locutrices). L'écart type est entre parenthèses. Les variations selon la syllabe les plus amples sont grisées	204
Tableau 68 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur le F1 des voyelles du français ...	204
Tableau 69 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur le F1 des voyelles du français.....	205
Tableau 70 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur le F2 des voyelles du français ...	206
Tableau 71 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur le F2 des voyelles du français.....	206
Tableau 72 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur le F3 des voyelles du français ...	207
Tableau 73 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur le F3 des voyelles du français.....	207
Tableau 74 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur la durée des voyelles tchèques ..	209
Tableau 75 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur la durée des voyelles du tchèque.....	209
Tableau 76 : Valeurs moyennes de F1 et F2 (toutes les voyelles), F3 (voyelles antérieures) des voyelles du tchèque, prononcées en syllabes initiale, médiane et finale dans des logatomes CVCVCVC où C = [p, t, k, ɦ] (valeur prise à la moitié de la durée vocalique* 4 répétitions* 4 contextes confondus* 20 locutrices). Ecart type est entre parenthèses. Les variations selon la syllabe les plus amples sont grisées	211
Tableau 77 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur le F1 des voyelles tchèques.....	211
Tableau 78 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur le F1 des voyelles du tchèque	212
Tableau 79 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur le F2 des voyelles tchèques.....	212
Tableau 80 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur le F2 des voyelles du tchèque	213
Tableau 81 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet de la syllabe sur le F3 des voyelles tchèques.....	214
Tableau 82 : Résultat de l'ANOVA à 2 facteurs mesurant l'effet du facteur Voyelle, Syllabe et Voyelle*Syllabe sur le F3 des voyelles du tchèque	214
Tableau 83 : Corpus de (Delattre, 1969) pour l'étude de la réduction des voyelles du français en position inaccentuée. La syllabe accentuée est précédée d'apostrophe. La voyelle dont le timbre est étudié est en italique. La voyelle finale imposant son lieu d'articulation est en gras	216
Tableau 84 : Informations sur la construction et passation du test de perception interlangue des voyelles isolées ..	218
Tableau 85 : Identification interlangue des voyelles du français par des Tchèques monolingues. Les résultats sont classés selon le <i>fit index</i> , calculé selon (Guion <i>et al.</i> , 2000). AC = aucune catégorie	220
Tableau 86 : Informations sur la construction et passation du test de perception interlangue des voyelles françaises en contextes pVp, tVt, kVk	222
Tableau 87 : Identification interlangue des voyelles du français en contexte labial (4 répétitions) par 10 Tchèques monolingues. Les résultats sont classés selon le <i>fit index</i> , calculé selon (Guion <i>et al.</i> , 2000). AC = aucune catégorie.....	224

Tableau 88 : Identification interlangue des voyelles du français en contexte dental (4 répétitions) par 10 Tchèques monolingues. Les résultats sont classés selon le <i>fit index</i> , calculé selon (Guion <i>et al.</i> , 2000). AC = aucune catégorie.....	226
Tableau 89 : Identification interlangue des voyelles du français en contexte palato-vélaire (4 répétitions) par 10 Tchèques monolingues. Les résultats sont classés selon le <i>fit index</i> , calculé selon (Guion <i>et al.</i> , 2000). AC = aucune catégorie	228
Tableau 90 : Similarité acoustique (exprimée par le nombre d'écarts types ET des formants F1/F2, (F2/F3)) et perceptive (exprimée par le <i>fit index</i>) entre les voyelles du français (FR) et du tchèque (TCH) en contextes nul (0), labial (p), dental (t) et palato-vélaire (k). Voyelles <i>nouvelles</i> et voyelles "hautement similaires" grisées (gris foncé et gris clair respectivement).....	231
Tableau 91 : Description du corpus et du calcul des moyennes formantiques des voyelles françaises produites en isolation et en contexte par des natives du français (la « référence ») et les non-natives.....	240
Tableau 92 : Information sur la construction et les conditions de passation du test d'identification des voyelles isolées du français produites par dix tchéophones et identifiées par dix Français natifs non-méridionaux...	241
Tableau 93 : Information sur la construction et les conditions de passation du test d'identification des voyelles isolées du français produites par une Française parisienne et identifiées par dix Tchèques, futures enseignantes de FLE.....	242
Tableau 94 : Identification par 10 Tchèques, futures enseignantes de FLE, du [ø] produit en isolation par une Française parisienne, répétées 8 fois. Le stimulus est en colonne 1, les réponses en colonne 2, le pourcentage de l'identification en colonne 3, la note moyenne en colonne 4 et le <i>fit index</i> , calculé selon Guion <i>et al.</i> (2000), en colonne 5	243
Tableau 95 : Moyennes et écarts types de F1, F2 et F3 (en Hertz) de [y] en contexte tVt et de [ø] en isolation (0) et en contexte kVk, produits par 10 Françaises natives (Nat) et 10 Tchèques (Non-nat), 4 répétitions. Les cases avec les valeurs à plus d'un écart type de la « référence » sont grisées	243
Tableau 96 : Le F2 moyen (en Hz) de [ø] dans [køk] chez les 10 non-natives (non-nat) séparément (moyenne de 4 répétitions).....	245
Tableau 97 : Similarité "symbolique" (S), acoustique (A) et perceptive (P) des dix voyelles françaises avec les voyelles tchèques selon le contexte phonétique. Divergence "symbolique" marquée entre parenthèses : malgré l'existence du symbole [e, o] en tchèque, le [e, o] français se réalise et/ou est identifié comme respectivement [i, u/u:] tchèques	246
Tableau 98 : Distance acoustique F4-F3 du [i] prononcé par 10 Françaises natives (Nat) et 10 Tchèques (Non-nat) en isolation (0) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], 4 répétitions par locutrice.....	248
Tableau 99 : Identification par 10 Tchèques, futures enseignantes de FLE, du [ɛ] produit en isolation par une Française parisienne, répétées 8 fois. Le stimulus est en colonne 1, les réponses en colonne 2, le pourcentage de l'identification en colonne 3, la note moyenne en colonne 4 et le <i>fit index</i> , calculé selon Guion <i>et al.</i> (2000), en colonne 5	249
Tableau 100 : Distance acoustique F2-F1 du [u] prononcé par 10 Françaises natives (Nat) et 10 Tchèques (Non-nat) en isolation (0) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], 4 répétitions par locutrice.....	251
Tableau 101 : Production et perception par 10 Tchèques des voyelles françaises [i, ɛ, u] du groupe 1 se caractérisant par une similarité "symbolique" (S), acoustique (A) et perceptive (P) avec les voyelles tchèques. Le niveau en production est estimé selon la distance des formants par rapport à la "référence" (exprimée en nombre d'écarts types) et le <i>fit index</i> de l'identification par 10 Français natifs de chaque voyelle produite par 10 Tchèques en isolation (100 réponses par voyelle). Le niveau de perception est estimé par le <i>fit index</i> de l'identification par 10 Tchèques de chacune des voyelles répétées 8 fois par une Française parisienne (80 réponses par voyelle)	252
Tableau 102 : Identification par 10 Tchèques, futures enseignantes de FLE, du [e] produit en isolation par une Française parisienne, répétées 8 fois. Le stimulus est en colonne 1, les réponses en colonne 2, le pourcentage de l'identification en colonne 3, la note moyenne en colonne 4 et le <i>fit index</i> , calculé selon Guion <i>et al.</i> (2000), en colonne 5	253
Tableau 103 : Distance acoustique F2-F1 du [o] prononcé par 10 Françaises natives (Nat) et 10 non-natives de LM tchèque (Non-nat) en isolation (0) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], 4 répétitions par locutrice	256
Tableau 104 : Production et perception par 10 Tchèques des voyelles françaises [e, o, a] du groupe 2 se caractérisant par une similarité acoustique (A) et perceptive (P) avec les voyelles tchèques pouvant être retranscrites selon le	

contexte avec un symbole différent. Le niveau en production est estimé selon la distance des formants par rapport à la “référence” (exprimée en nombre d’écarts types) et le <i>fit index</i> de l’identification par 10 Français natifs de chaque voyelle produite par 10 Tchèques (100 réponses par voyelle). Le niveau de perception est estimé par le <i>fit index</i> de l’identification par 10 Tchèques de chacune des voyelles répétées 8 fois par une Française parisienne (80 réponses par voyelle)	258
Tableau 105 : Identification par 10 Tchèques, futures enseignantes de FLE, du [ɔ] produit en isolation par une Française parisienne, répétées 8 fois. Le stimulus est en colonne 1, les réponses en colonne 2, le pourcentage de l’identification en colonne 3, la note moyenne en colonne 4 et le <i>fit index</i> , calculé selon Guion <i>et al.</i> (2000), en colonne 5	259
Tableau 106 : Distance acoustique F2-F1 (en Hz) du [ɔ] prononcé par 10 Françaises natives (Nat) et 10 Non-natives (Non-nat) en isolation (0) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], 4 répétitions par locutrice ...	260
Tableau 107 : Production et perception par 10 Tchèques du [ɔ] se caractérisant par une similarité perceptive (P) et acoustique restreinte (A) avec les voyelles tchèques pouvant être retranscrites selon le contexte avec un symbole différent. Le niveau en production est estimé selon la distance des formants par rapport à la “référence” (exprimée en nombre d’écarts types) et le <i>fit index</i> de l’identification par 10 Français natifs de chaque voyelle produite par 10 Tchèques (100 réponses par voyelle). Le niveau de perception est estimé par le <i>fit index</i> de l’identification par 10 Tchèques de chacune des voyelles répétées 8 fois par une Française parisienne (80 réponses par voyelle)	261
Tableau 108 : Identification par 10 Tchèques, futures enseignantes de FLE, du [œ] produit en isolation par une Française parisienne, répétées 8 fois. Le stimulus est en colonne 1, les réponses en colonne 2, le pourcentage de l’identification en colonne 3, la note moyenne en colonne 4 et le <i>fit index</i> , calculé selon Guion <i>et al.</i> (2000), en colonne 5	262
Tableau 109 : Production et perception par 10 Tchèques du [œ] se caractérisant par un symbole <i>nouveau</i> et une similarité perceptive (P) et acoustique (A) limitée aux contextes labial et coronal. Le niveau en production est estimé selon la distance des formants par rapport à la “référence” (exprimée en nombre d’écarts types) et le <i>fit index</i> de l’identification par 10 Français natifs de chaque voyelle produite par 10 Tchèques (100 réponses par voyelle). Le niveau de perception est estimé par le <i>fit index</i> de l’identification par 10 Tchèques de chacune des voyelles répétées 8 fois par une Française parisienne (80 réponses par voyelle)	263
Tableau 110 : Les voyelles françaises <i>nouvelles</i> [y, ø] selon le contexte (nul (0) ou [p,t,k]) du point de vue “symbolique” (S), acoustique (A) et perceptif (P). Les voyelles tchèques perceptivement <i>similaires</i> sont entre parenthèses. Le contexte qui rend la voyelle <i>nouvelle</i> selon les trois critères est grisé	264
Tableau 111 : Distance acoustique F3-F2 du [y] prononcé par 10 Françaises natives (Nat) et 10 non-natives (Non-nat) en isolation (0) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], 4 répétitions par locutrice	265
Tableau 112 : Production et perception par 10 Tchèques des [y, ø] se caractérisant par un symbole <i>nouveau</i> et, selon le contexte, une similarité perceptive (P). Le niveau en production est estimé selon la distance des formants par rapport à la “référence” (exprimée en nombre d’écarts types) et le <i>fit index</i> de l’identification par 10 Français natifs de chaque voyelle produite par 10 Tchèques (100 réponses par voyelle). Le niveau de perception est estimé par le <i>fit index</i> de l’identification par 10 Tchèques de chacune des voyelles répétées 8 fois par une Française parisienne (80 réponses par voyelle). Les contextes où ces voyelles sont <i>similaires</i> sont grisés	267
Tableau 113 : Voyelles du français non maîtrisées par les futures enseignantes de FLE	269
Tableau 114 : Identification par 10 Français natifs des 10 voyelles orales produites en isolation par 10 futures enseignantes de FLE. Les stimuli sont en colonne 1, les réponses en colonne 2, le pourcentage de l’identification en colonne 3, la note moyenne en colonne 4 et le <i>fit index</i> , calculé selon Guion <i>et al.</i> (2000), en colonne 5	271
Tableau 115 : Identification par 10 Tchèques, futures enseignantes de FLE, des 10 voyelles orales produites en isolation par une Française parisienne, répétées 8 fois. Les stimuli sont en colonne 1, les réponses en colonne 2, le pourcentage de l’identification en colonne 3, la note moyenne en colonne 4 et le <i>fit index</i> , calculé selon Guion <i>et al.</i> (2000), en colonne 5	272
Tableau 116 : Informations sur les 5 natifs du français ayant répondu au questionnaire phonologique et effectué les tests de perception sur les voyelles moyennes du français	277
Tableau 117 : Information sur la construction et les conditions de passation du test de discrimination des voyelles moyennes du français par 5 Français natifs non méridionaux	278
Tableau 118 : Suites de mots proposées aux 5 auditeurs Français natifs non méridionaux dans la tâche de la singularité. Les suites grisées sont des stimuli d’entraînement ou de contrôle	279

Tableau 119 : Résultat de la tâche de la singularité sur la discrimination des voyelles moyennes dans des suites de 4 mots, effectuée par 5 auditeurs français natifs non méridionaux (leur région d'origine est indiquée en colonne 1).....	280
Tableau 120 : Information sur la construction et les conditions de passation du test d'identification des voyelles moyennes du français par 5 Français natifs non méridionaux	280
Tableau 121 : 24 phrases proposées aux 5 auditeurs Français natifs non méridionaux qui devaient identifier les phrases prononcées avec un accent régional pouvant être induit par l'inversion du timbre au sein des paires de voyelles moyennes (mots en gras). La phrase où l'inversion n'a pas été identifiée par l'auditeur 2 est soulignée	281
Tableau 122 : Résultat de la tâche d'identification par 5 Français non-méridionaux de 4 phrases prononcées avec [e] en syllabe CVC, 4 avec [ɔ] et 4 avec [œ] en syllabe CV	282
Tableau 123 : Information sur la construction et les conditions de passation du test d'identification par 5 Français natifs non méridionaux des mots dont le contraste est basé sur la qualité des voyelles moyennes du français	282
Tableau 124 : Paires minimales proposées dans le test d'identification des mots contenant une voyelle moyenne. La qualité effectivement prononcée est en gras.....	283
Tableau 125 : Résultat du test d'identification par 5 Français natifs non méridionaux de 26 mots basés sur le contraste e/ɛ, 31 mots sur o/ɔ et 4 mots sur ø/œ	284
Tableau 126 : Liste des paires minimales dans lesquelles les contrastes e/ɛ, o/ɔ et ø/œ sont systématiquement perçus par 5 auditeurs français natifs non-méridionaux	284
Tableau 127 : Information sur la construction et les conditions de passation du test d'identification des voyelles moyennes du français par 5 Français natifs non méridionaux	285
Tableau 128 : 58 phrases proposées aux 5 auditeurs Français natifs non méridionaux qui ont identifié les phrases ayant un sens inattendu (bizarre, drôle) induit par l'inversion du timbre au sein des paires des voyelles moyennes (mots en gras). Les phrases dont le sens est inattendu selon tous les 5 auditeurs sont soulignées ...	287
Tableau 129 : Résultat de la tâche d'identification de 31 phrases bizarres effectuée par 5 auditeurs français natifs non méridionaux	287
Tableau 130 : Construction et passation du test d'identification des sons vocaliques synthétisées avec Klatt	289
Tableau 131 : Identification des voyelles synthétisées en termes de /e/ ou /ɛ/ (en italique) avec la valeur du <i>fit index</i> de l'identification, calculé sur la base de la proportion de l'identification et de la note moyenne obtenue.....	291
Tableau 132 : Identification des voyelles synthétisées en termes de /o/ ou /ɔ/ (en italique) avec la valeur du <i>fit index</i> de l'identification, calculé sur la base de la proportion de l'identification et de la note moyenne obtenue.....	291
Tableau 133 : Identification des voyelles synthétisées en termes de /ø/ ou /œ/ (en italique) avec la valeur du <i>fit index</i> de l'identification, calculé sur la base de la proportion de l'identification et de la note moyenne obtenue.....	292
Tableau 134 : Nombre de paires minimales basées sur le contraste des voyelles moyennes qui ne sont jamais prononcées de la même façon selon 5 enquêtés français natifs non méridionaux	297
Tableau 135 : La même réalisation des mots de 14 paires minimales dont le contraste repose sur les voyelles moyennes e/ɛ. Questionnaire rempli par 5 Français natifs non-méridionaux ; le no de l'auditeur ayant choisi la réponse indiquée se trouve entre parenthèses.....	297
Tableau 136 : Réalisation identique des mots de 14 paires minimales dont le contraste repose sur les voyelles moyennes o/ɔ. Questionnaire rempli par 5 Français natifs non-méridionaux : le no de l'auditeur ayant choisi la réponse indiquée se trouve entre parenthèses.....	298
Tableau 137 : Réalisation identique des mots de 2 paires minimales dont le contraste repose sur les voyelles moyennes ø/œ. Questionnaire rempli par 5 Français natifs non-méridionaux : le no de l'auditeur ayant choisi la réponse indiquée se trouve entre parenthèses.....	298
Tableau 138 : Les valeurs formantiques F1, F2 et F3 des voyelles [e] et [ɛ] produits par 10 Françaises natives (4 répétitions) en isolation et en syllabe finale pVp, tVt, kVk, RVR (valeur prise à la moitié de la durée du segment), en Hertz (1 ^e colonne), en Bark (2 ^e colonne) et la différence en pourcentage (3 ^e colonne)	300
Tableau 139 : Les valeurs formantiques F1, F2 et F3 des voyelles [e] et [ɛ] produits par une Française native en isolation et en syllabe finale pVp, tVt, kVk, RVR (valeur prise à la moitié de la durée du segment)	301
Tableau 140 : Les valeurs formantiques F1 et F2 des voyelles [o] et [ɔ] produits par 10 Françaises natives (4 répétitions) en isolation et en syllabe finale pVp, tVt, kVk, RVR (valeur prise à la moitié de la durée du	

segment), en Hertz (1 ^e colonne), en Bark (2 ^e colonne), la différence en pourcentage (3 ^e colonne) et la distance F2-F1 (4 ^e colonne)	302
Tableau 141 : Les valeurs formantiques F1 et F2 (en Hertz) des voyelles [o] et [ɔ] produits par une Française native en isolation et en syllabe finale pVp, tVt, kVk, RVR (valeur prise à la moitié de la durée du segment)	303
Tableau 142 : Les valeurs formantiques F1, F2 et F3 des voyelles [ø] et [œ] produites par 10 Françaises natives (4 répétitions) en isolation et en syllabe finale pVp, tVt, kVk, RVR (valeur prise à la moitié de la durée du segment), en Hertz (1 ^e colonne), en Bark (2 ^e colonne) et la différence en pourcentage (3 ^e colonne)	305
Tableau 143 : Valeurs moyennes de F1, F2 et F3 (en Hz) de [ø] et [œ] produits par une Française native en isolation et en syllabe finale pVp, tVt, kVk, RVR (valeur prise à la moitié de la durée du segment).....	306
Tableau 144 : Résultat de la tâche de la singularité sur la discrimination des voyelles moyennes dans des suites de 4 monosyllabes, effectuée par 10 auditrices tchèques, futures enseignantes de FLE. Les meilleurs résultats de discrimination des deux timbres sont en gras	310
Tableau 145 : Corpus sur les voyelles moyennes du français et calcul des moyennes formantiques.....	311
Tableau 146 : Le <i>fit index</i> de l'identification des voyelles moyennes du français produites par les locutrices tchèques T1 à T10 (colonne 1), 10 répétitions par voyelle, perçues par 10 Français natifs. La valeur du <i>fit index</i> va de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte). Les voyelles produites avec authenticité (avec <i>fit index</i> supérieur à 4) sont griséées.....	312
Tableau 147 : Le <i>fit index</i> de l'identification du [ø] en termes de /ø/ ou /œ/ (ou /y/). Les voyelles ont été produites par les locutrices tchèques T1 à T10, 10 répétitions, identifiées par 10 Français natifs. Le <i>fit index</i> s'étend de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte). Le <i>fit index</i> égal ou supérieur à 4 (reflétant l'atteinte authentique de la cible) est en gras	313
Tableau 148 : Valeurs de F1, F2, F3 (en Hz) de [ø] isolé produits par les locutrices tchèques 1, 2, 4, 5, 6, 7, 8, 9 et 10. Valeurs relevées à la moitié de la durée vocalique. La valeur minimale et maximale pour chaque formant est en gras.....	314
Tableau 149 : Valeurs de F1, F2, F3 (en Hz) de 10 occurrences du [ø] isolé produit par la locutrice tchèque T3 et identifiées par 10 Français natifs. Valeurs relevées à la moitié de la durée vocalique. La valeur minimale et maximale pour chaque formant est en gras	315
Tableau 150 : Le <i>fit index</i> de l'identification du [ɛ] en termes de /ɛ/, /e/ ou /æ/. Les voyelles ont été produites par les locutrices tchèques T1 à T10, 10 répétitions, identifiées par 10 Français natifs. Le <i>fit index</i> s'étend de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte). Le <i>fit index</i> égal ou supérieur à 4 (reflétant l'atteinte authentique de la cible) est en gras	317
Tableau 151 : Valeurs de F1, F2, F3 (en Hz) de [ɛ] isolé produits par les locutrices tchèques 2, 3, 4, 5, 6, 9 et 10. Valeurs relevées à la moitié de la durée vocalique. La valeur minimale et maximale pour chaque formant est en gras.....	317
Tableau 152 : Le <i>fit index</i> de l'identification du [e] en termes de /e/ ou /ɛ/. Les voyelles ont été produites par les locutrices tchèques T1 à T10, 10 répétitions, identifiées par 10 Français natifs. Le <i>fit index</i> s'étend de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte). Le <i>fit index</i> égal ou supérieur à 4 (reflétant l'atteinte authentique de la cible) est en gras	318
Tableau 153 : Le <i>fit index</i> de l'identification du [o] en termes de [o] ou [ɔ]. Les voyelles ont été produites par les locutrices tchèques T1 à T10, 10 répétitions, identifiées par 10 Français natifs. Le <i>fit index</i> s'étend de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte). Le <i>fit index</i> égal ou supérieur à 4 (reflétant l'atteinte authentique de la cible) est en gras	321
Tableau 154 : Le <i>fit index</i> de l'identification du [œ] en termes de /œ/ ou /ø/. Les voyelles ont été produites par les locutrices tchèques T1 à T10, 10 répétitions, identifiées par 10 Français natifs. Le <i>fit index</i> s'étend de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte). Notons que le [œ] des seules locutrices T3 et T4 a été majoritairement catégorisé comme tel et peu confondu avec /ø/	323
Tableau 155 : Valeurs de F1, F2, F3 (en Hz) et identification par 10 auditeurs français natifs de [œ] isolé produit par la locutrice tchèque T3 (10 répétitions). Valeurs relevées à la moitié de la durée vocalique. La valeur minimale et maximale pour chaque formant est en gras.....	324
Tableau 156 : Valeurs de F1, F2, F3 (en Hz) et identification par 10 auditeurs français natifs de [œ] isolé produit par la locutrice T4 (10 répétitions). Valeurs relevées à la moitié de la durée vocalique. La valeur minimale et maximale pour chaque formant est en gras	326

Tableau 157 : Le <i>fit index</i> de l'identification du [ɔ] en termes de /ɔ/ ou /o/. Les voyelles ont été produites par les locutrices tchèques T1 à T10, 10 répétitions, perçues par 10 Français natifs. Le <i>fit index</i> s'étend de 0 à 5 (0 = cible vocalique non atteinte, 5 = cible atteinte)	328
Tableau 158 : Valeurs de F1 et F2 (en Hz) et identification par 10 auditeurs français natifs du [ɔ] isolé produit par la locutrice tchèque T9 (10 répétitions). Valeurs relevées à la moitié de la durée vocalique. La valeur minimale et maximale pour chaque formant est en gras	329

Liste des figures

Figure 1 : Structure de la thèse	20
Figure 2 : Formation des sons de la parole selon la théorie source-filtre.	26
Figure 3 : Modèle avec un tube pour [œ], avec la glotte à gauche et les lèvres à droite. Adaptation selon (Fant 1960)	28
Figure 4 : Modèle de deux tubes pour [a] avec la glotte à gauche et les lèvres à droite. 1 = cavité buccale, 2 = cavité pharyngale. Adaptation selon Fant (1960)	28
Figure 5 : Modèle de deux tubes pour [i] (à gauche) et [y] (au milieu), avec la glotte à gauche et les lèvres à droite. 1 = cavité buccale, 2 = cavité pharyngale. Adaptation selon (Fant 1960). Changement d'affiliation de cavités que l'on peut voir sur le spectrogramme du [i] et [y] (à droite)	29
Figure 6 : Modèle à quatre tubes et nomogrammes de Fant (1960) représentant la glotte à droite, les lèvres à gauche. Les flèches indiquent les points focaux	30
Figure 7 : Les points de convergences formantiques dans un nomogramme (image du haut) et spectrogramme correspondant des 6 voyelles focales du français [u, o, ɔ, a, i, y], prononcées par un locuteur Français natif, selon (Vaissière, 2011)	31
Figure 8 : Effet des contextes consonantiques labial, post-dental et vélaire sur les formants F1 et F2 de huit voyelles de l'anglais. Selon Stevens and House (1963).....	35
Figure 9 : Influence du contexte consonantique sur les formants F1/F2 (en Hz) des voyelles anglaises prononcées par 6 hommes dont les valeurs sont relevées à la moitié (image du haut) et au début ou à la fin (image du bas, respectivement à gauche et à droite) de la durée vocalique. Selon Hillenbrand <i>et al.</i> (2001).....	36
Figure 10 : F1 et F2 moyens (en Hz) des voyelles du français [i, e, ε, a, α, u, o, ɔ, y, ø, œ] (4 locuteurs français natifs*11 voyelles* 2 répétitions) en isolation (rond blanc), en syllabe tVt (carré), en syllabe tV (rond noir) et en syllabe Vt (étoile), selon Gottfried (1984)	37
Figure 11 : Triangle vocalique F1/F2 (en Bark) des voyelles du français (en gras) et de l'anglais, produites en contextes bilabiale (à gauche) et alvéolaire (à droite) par 3 locuteurs natifs selon Levy (2009)	43
Figure 12 : Les enveloppes spectrales des sons de base utilisés pour tester le « difference limen » de F1, selon Flanagan (1955).....	45
Figure 13 : Les enveloppes spectrales des sons de base utilisés pour tester le « difference limen » de F2, selon Flanagan (1955).....	46
Figure 14 : Le pourcentage moyen d'erreurs d'identification des voyelles du français selon le contexte syllabique auprès des auditeurs français natifs (en noir) et des anglophones américains parlant français (en blanc), selon Gottfried (1984).....	48
Figure 15 : Le pourcentage moyen d'erreurs de discrimination des voyelles du français selon le contexte syllabique auprès des auditeurs français natifs (en noir) et des anglophones américains parlant français (en blanc) et des anglophones américains monolingues (avec rayures), selon Gottfried (1984)	48
Figure 16 : Le triangle des voyelles monophthongues du tchèque de Bohême (à gauche) et de Moravie (à droite), selon Simackova <i>et al.</i> (2012).....	51
Figure 17 : 11 voyelles orales du français (entourées), selon Vaissière (2009)	52
Figure 18 : Comparaison des triangles vocaliques sur le plan F1-F2 (en Bark) à partir de données de Gendrot and Adda-Decker (2005) : moyennes de productions de 15 femmes, 2 heures de parole radiophonique (en pointillé), et du Groupe Didactique (Georgeton <i>et al.</i> , 2012) : moyennes de productions de 40 femmes, 4 répétitions, voyelles isolées (en trait plein). Les ellipses sont tracées à un écart type des moyennes F1-F2	56
Figure 19 : Triangle F2/F3 (en Bark) des voyelles [i, e, ε, y, ø, œ, a] en isolation (40 femmes, 4 répétitions par voyelle), selon les données du GD (Georgeton <i>et al.</i> , 2012). Les ellipses sont tracées à 1 écart type des moyennes F2-F3	57
Figure 20 : Triangle vocalique F1/F2 (en Hz) des voyelles tchèques à partir de productions de 4 locuteurs natifs, selon Hála (1941).....	57
Figure 21 : Triangle F1/F2 (en Hertz) des voyelles brèves du tchèque [i, ε, u, a, o] produites en isolation par 21 hommes et 14 femmes, selon Grepl <i>et al.</i> (2007)	58

Figure 22 : Triangle F1/F2 (en ERB-Hz) des voyelles du tchèque [ɪ, i:, ɛ, ɛ:, a, a:, o, o:, u, u:] issue de la tâche de lecture par 27 hommes (à gauche) et 48 femmes (à droite), selon Skarnitzl and Volin (2012).....	58
Figure 23 : Complexité des structures syllabiques en allemand, anglais, tchèque, français et italien, selon Dankovicova and Dellwo (2007)	65
Figure 24 : Relation entre les phonèmes de la LM et de la LE: 2 catégories néerlandaises natives /u, ɔ/ correspondant à une catégorie espagnole non native /u/ (à gauche) ou une catégorie japonaise native /r/ correspondant à 2 catégories anglaises non-natives /l, ɹ/ (split – scission - allophonique, à droite), selon (Escudero and Boersma, 2001)	69
Figure 25 : Modélisation de [y] par Ouni and Laprie (2001)	72
Figure 26 : Modélisation de [u] par Ouni and Laprie (2001).....	73
Figure 27 : Apprenants japonophones : niveau de difficulté de l'apprentissage phonétique des voyelles françaises [a, ø, y, u] selon leur similarité phonémique/ phonétique avec les voyelles du japonais selon Kamiyama and Vaissière (2009).....	79
Figure 28 : Corrélations entre triangle acoustique (image 1) et triangles articulatoires (image 2 à 4), selon Crothers (1978).....	89
Figure 29 : Interface graphique du logiciel VTDemo (Huckvale, 2009)	90
Figure 30 : Interface graphique du logiciel WaveSurfer avec spectrogramme des voyelles [i, y] produites par un Français natif	91
Figure 31 : Locutrices ayant produit le corpus français (10 francophones natives et 10 tchécophones natives - futures enseignantes de FLE de LM tchèque) et le corpus tchèque (10 tchécophones natives - futures enseignantes de FLE de LM tchèque et 10 tchécophones natives non parlant le français).....	105
Figure 32 : Phase d'entraînement des voyelles du français insérées en contexte phonétique.....	111
Figure 33 : Spectrogramme des voyelles françaises [u, o, ɔ], prononcées par une Française native en isolation (iso) et en contexte symétrique dental (t)	114
Figure 34 : Fonctionnalités générales de la partie Client du logiciel VisuVo	116
Figure 35 : Fonctionnalités propres au graphique « Chart » de la partie Client du logiciel VisuVo.....	117
Figure 36 : Fonctionnalités propres au graphique « Chart » des voyelles isolées de la partie Client du logiciel VisuVo	118
Figure 37 : Triangle vocalique du français standard actuel (à gauche) et du tchèque de Bohême (à droite), les symboles divergents sont entre parenthèses.....	125
Figure 38 : Triangle vocalique des voyelles orales isolées du français (en pointillé) et du tchèque (en trait plein) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 10 Françaises/ 20 Tchèques*4 répétitions*3 valeurs par voyelle.....	132
Figure 39 : Spectrogramme des voyelles orales isolées [i, e, ɛ, ɘ, ɔ, o, u, œ, ø, y] prononcées par une Française native	133
Figure 40 : Spectrogramme des voyelles tchèques [i:, ɪ, ɛ, ɛ:, ɘ, a:, i, o:, u, u:] prononcées par une Tchèque native	133
Figure 41 : Spectrogramme des voyelles [u, ɘ, i] du français et [u:, a:, i:] du tchèque, prononcées isolément par une Française et une Tchèque native	134
Figure 42 : Spectrogramme des voyelles du français [u, o, ɔ, ɘ, y, i] prononcées en isolation par une Française native	136
Figure 43 : Spectrogramme des voyelles du tchèque [u, u:, o, o:, ɘ, a:, ɪ, i:] prononcées en isolation par une Tchèque native.....	136
Figure 44 : Les formants F1 – F4 moyens (en Bark) de la classe des voyelles fermées du français [i, y, u]. L'écart type est de 1	137
Figure 45 : Les formants F1 – F3 moyens (en Bark) de la classe des voyelles mi-fermées du français [e, ø, o]. L'écart type est de 1	138
Figure 46 : Les formants F1 – F3 moyens (en Bark) de la classe des voyelles mi-ouvertes du français [ɛ, œ, ɔ]. L'écart type est de 1	139

Figure 47 : Les formants F1 – F3 moyens (en Bark) des voyelles postérieures du français [u, o, ɔ] et du [a] avec une réalisation pharyngale en isolation. L'écart type est de 1.....	140
Figure 48 : Les formants F1 – F4 moyens (en Bark) des voyelles fermées du tchèque [i:, ɪ, u:, u]. L'écart type est de 1	141
Figure 49 : Les formants F1 – F3 moyens (en Bark) des voyelles moyennes du tchèque [ɛ:, ɛ, o:, o]. L'écart type est de 1	142
Figure 50 : Les formants F1 – F3 moyens (en Bark) des voyelles postérieures du tchèque [u:, u, o, o:] et des voyelles centrales [a, a:]. L'écart type est de 1	142
Figure 51 : Formants moyens (en Bark) des voyelles françaises antérieures (en haut) et postérieures (en bas) qui se distinguent essentiellement par la valeur du F2. L'écart type est de 1.....	143
Figure 52 : Formants moyens (en Bark) des voyelles tchèques antérieures (en haut) et postérieures (en bas) qui se distinguent essentiellement par la valeur du F2. L'écart type est de 1.....	144
Figure 53 : Formants moyens (en Bark) des voyelles françaises antérieures labiales et étirées, ayant F2 et F3 plus élevés. L'écart type est de 1.....	144
Figure 54 : Durée moyenne (en ms) des dix voyelles orales tchèques isolées [ɛ, ɛ:, a, a:, ɪ, i:, o, o:, u, u:] chacune produite 4 fois par 20 Tchèques natives	145
Figure 55 : Formants moyens (en Bark) des voyelles tchèques brèves [ɛ, o, a] et de leurs homologues longues [ɛ:, o:, a:]. L'écart type est de 1	146
Figure 56 : Formants moyens (en Bark) des voyelles tchèques brèves [ɪ, u] et de leurs homologues longues [i:, u:]. L'écart type est de 1.....	146
Figure 57 : Triangle vocalique (en Bark) des voyelles orales isolées du français (en pointillé) et du tchèque (en trait plein) sur le plan F1/F2 (en haut) et F2/F3 (en bas). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 10 Françaises/ 20 Tchèques*4 répétitions*3 valeurs par voyelle	150
Figure 58 : Spectrogramme des paires de voyelles du français et du tchèque [e-ɪ], [a-a], [i-i:], [ɔ-o], [u-u:], [ɛ-ɛ], [o-u] prononcées par une Française et une Tchèque. Notons la similarité des patrons formantiques	152
Figure 59 : Valeurs de F1 et F2 (en Bark) de [ɔ] produit en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ɾ], calculées à partir de 4 répétitions par 10 Françaises natives et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1	158
Figure 60 : Spectrogramme des voyelles [i, y] du français prononcées isolément et en contexte RVR. Notons que le contexte uvulaire engendre une élévation du F1 et un abaissement des formants principalement dus à la cavité antérieure (F3 pour [i] et F2 pour [y]).	159
Figure 61 : Le F1 des dix voyelles orales du français [i, e, ɛ, y, ø, œ, u, o, ɔ, a] prononcées isolément (0) et en contextes [p, t, k, ɾ] dans des logatomes CVCVCVC par 10 locutrices*4 répétitions.	159
Figure 62 : Le F2 des dix voyelles orales du français [i, e, ɛ, y, ø, œ, u, o, ɔ, a] prononcées isolément (0) et en contextes [p, t, k, ɾ] dans des logatomes CVCVCVC (10 locutrices*4 répétitions)	162
Figure 63 : Spectrogramme des voyelles [y, ø, œ, i] du français prononcées par une Française isolément et en contextes pVp et/ou kVk qui provoquent par rapport à la cible des mouvements minimes, comme expliqué dans le paragraphe ci-dessus et indiqué dans le Tableau 39	163
Figure 64 : Triangle vocalique des voyelles orales du français en isolation (trait plein) et en contexte symétrique labial (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 10 Françaises *4 répétitions de chaque voyelle*valeur prise à la moitié de la durée vocalique	164
Figure 65 : Spectrogramme des voyelles [u, o, ɔ, a] du français prononcées isolément et en contexte pVp par une Française native. Notons que le contexte labial engendre une augmentation du F2 des voyelles postérieures de même que du [a] qui est pharyngal en isolation	165
Figure 66 : Triangle vocalique des voyelles orales du français en isolation (trait plein) et en contexte symétrique dental (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 10 Françaises *4 répétitions de chaque voyelle*valeur prise à la moitié de la durée vocalique	166
Figure 67 : Spectrogramme des voyelles [u, o, ɔ] du français prononcées isolément et en contexte tVt. Notons que le contexte labial engendre une augmentation importante du F2 des voyelles postérieures du français. Notons que	

le caractère focal des voyelles postérieures, qui rapprochent les formants F2-F1 en contexte nul, est particulièrement affaibli en contexte [t]	166
Figure 68 : Triangle vocalique des voyelles orales du français en isolation (trait plein) et en contexte symétrique palato-vélaire (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 10 Françaises *4 répétitions de chaque voyelle*valeur prise à la moitié de la durée vocalique	167
Figure 69 : Triangle vocalique des voyelles orales du français en isolation (trait plein) et en contexte symétrique uvulaire (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 10 Françaises *4 répétitions de chaque voyelle*valeur prise à la moitié de la durée vocalique	168
Figure 70 : Spectrogramme des voyelles [y, ø, e, ε] du français prononcées par une Française native isolément et en contexte RVR qui engendre une baisse du F2.....	168
Figure 71 : Spectrogramme des voyelles [u, o, ə] du français prononcées par une Française isolément et en contexte RVR qui ne provoque pas de mouvements importants du F2. Notons également que le caractère focal des voyelles postérieures, qui rapprochent les formants F2-F1 en contexte nul, est maintenu ou renforcé en contexte uvulaire.....	169
Figure 72 : Le F3 moyen des voyelles orales du français [i, e, ε, y, ø, œ, a] prononcées isolément (0) et en contextes [p, t, k, ʁ] dans des logatomes CVCVCVC (10 locutrices* 4 répétitions* valeur prise à la moitié de la durée vocalique).....	170
Figure 73 : Spectrogramme des voyelles [i, e, ε, y, ø, œ] du français prononcées par une Française isolément et en contexte RVR dans lequel le F3 des voyelles étirées baisse alors que celui des voyelles labiales augmente	172
Figure 74 : Spectrogramme de la voyelle [i] du français prononcée par une Française native isolément et en contextes pVp, tVt, kVk, RVR. Notons que seul le contexte uvulaire provoque un mouvement du F4 (qui baisse) par rapport à la cible	173
Figure 75 : Spectrogramme du [y] français prononcé par une Française native en contextes nul et [ʁ, k]. Notons que le caractère focal du [y], qui rapproche les formants F3-F2 en contextes nul et [k], est effacé en contexte [ʁ]	175
Figure 76 : Spectrogramme des voyelles [ε, a] du tchèque prononcées par une Tchèque isolément et en contexte kVk qui provoque une baisse de F1. Notons que le [a] en isolation est postérieur	180
Figure 77 : Le F1 moyen des dix voyelles monophthongues du tchèque [ɪ, i:, ε, ε:, o, o:, u, u:, a, a:] prononcées isolément (0) et en contextes [p, t, k, f] dans des logatomes CVCVCVC (20 locutrices*4 répétitions).....	180
Figure 78 : Le F2 moyen des dix voyelles monophthongues du tchèque [ɪ, i:, ε, ε:, o, o:, u, u:, a, a:] prononcées isolément (0) et en contextes [p, t, k, f] dans des logatomes CVCVCVC (20 locutrices*4 répétitions).....	182
Figure 79 : Triangle vocalique des voyelles monophthongues du tchèques en isolation (trait plein) et en contexte symétrique labial (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 20 Tchèques *4 répétitions de chaque voyelle*valeur prise à la moitié de la durée vocalique.....	183
Figure 80 : Triangle vocalique des voyelles monophthongues du tchèque en isolation (trait plein) et en contexte symétrique dental (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 20 Tchèques *4 répétitions de chaque voyelle*valeur prise à la moitié de la durée vocalique.....	184
Figure 81 : Triangle vocalique des voyelles monophthongues du tchèque en isolation (trait plein) et en contexte symétrique palato-vélaire (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 20 Tchèques *4 répétitions de chaque voyelle*valeur centrale.....	184
Figure 82 : Triangle vocalique des voyelles monophthongues du tchèque en isolation (trait plein) et en contexte symétrique glottal (en pointillé) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 1 écart type de la moyenne calculée à partir de productions de 20 Tchèques *4 répétitions de chaque voyelle*valeur prise à la moitié de la durée vocalique.....	185
Figure 83 : Le F3 moyen des voyelles monophthongues du tchèque [ɪ, i:, ε, ε:, a, a:] prononcées isolément (0) et en contextes [p, t, k, f] dans des logatomes CVCVCVC (20 locutrices*4 répétitions).....	186
Figure 84 : Spectrogramme de la voyelle [ɪ] du tchèque prononcée par une Tchèque isolément et en contextes pVp, tVt, kVk, hVh. Notons qu'aucun contexte n'engendre de modifications importantes du F4.....	188

Figure 85 : Les voyelles [a, a:, ε, ε:] du tchèque en contexte kVk prononcées par une Tchèque native. Notons un rapprochement des formants F2 et F3 sous forme d'une « pince vélaire ».....	191
Figure 86 : Triangles vocaliques F1/F2 (en Bark) des voyelles du français (image du haut) et des voyelles du tchèque (image du bas) en contextes nul (un rond, trait plein), labial (un triangle), dental (une croix), palato-vélaire (un losange) et uvulaire ou glottal (un triangle). Moyennes calculées à partir de productions de 10 Françaises et 20 Tchèques*4 répétitions*3 syllabes (valeurs prises à la moitié de la durée vocalique)	193
Figure 87 : Spectrogramme de la voyelle [ε] prononcée en isolation et en contexte kVk par une Française parisienne et une Tchèque de Bohême. Le F2 du [ε] en contexte kVk augmente par rapport au [ε] isolé davantage en tchèque qu'en français	194
Figure 88 : Les formants moyens F1, F2 et F3 (en Bark) de la voyelle [a] du français (représenté avec une croix) et du tchèque (représenté avec un rond) en isolation (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ] pour le français et [p, t, k, fi] pour le tchèque. Les moyennes sont calculées à partir de productions de 10 Françaises/20 Tchèques* 4 répétitions	195
Figure 89 : Les formants moyens F1, F2, F3 (en Bark) du [œ] français (représenté avec une croix, calculés à partir de 4 répétitions par 10 Françaises) et du [ε] tchèque (représenté avec un rond, calculés à partir de 4 répétitions par 20 Tchèques) en contexte nul et dans les logatomes CVCVCVC où C = [p, t, k, R/ h]. Les traits horizontaux représentent les formants des voyelles isolées. L'écart type est de 1.....	198
Figure 90 : Variation de la durée vocalique selon la position dans le mot : Syllabe 1 = initiale, Syllabe 2 = médiane, Syllabe 3 = finale	203
Figure 91 : F1 des voyelles orales du français en position initiale (en rouge), médiane (en vert) et finale (en bleu), calculé à partir de la valeur centrale*4 répétitions*4 contextes confondus*10 locutrices	205
Figure 92 : F2 des voyelles orales du français en positions initiale (en rouge), médiane (en vert) et finale (en bleu), calculé à partir de la valeur prise à la moitié de la durée vocalique* 4 répétitions* 4 contextes confondus* 10 locutrices.....	206
Figure 93 : F3 des voyelles orales du français en positions initiale (en rouge), médiane (en vert) et finale (en bleu), calculé à partir de la valeur prise à la moitié de la durée vocalique* 4 répétitions* 4 contextes confondus* 10 locutrices.....	207
Figure 94 : Variation de durée des dix voyelles orales du tchèque [i, i:, ε, ε:, o, o:, u, u:, a, a:] selon la position prosodique. Syllabe 1 = initiale, syllabe 2 = médiane, syllabe 3 = finale	209
Figure 95 : Variation de durée vocalique en tchèque selon la position prosodique. Syllabe 1 = initiale, syllabe 2 = médiane, syllabe 3 = finale	210
Figure 96 : F1 des voyelles monophthongues du tchèque en positions initiale (en rouge), médiane (en vert) et finale (en bleu), calculé à partir de la valeur centrale*4 répétitions*4 contextes confondus*20 locutrices.....	212
Figure 97 : F2 des voyelles monophthongues du tchèque en positions initiale (en rouge), médiane (en vert) et finale (en bleu), calculé à partir de la valeur prise à la moitié de la durée vocalique* 4 répétitions* 4 contextes confondus* 20 locutrices	213
Figure 98 : F3 des voyelles monophthongues du tchèque en positions initiale (en rouge), médiane (en vert) et finale (en bleu), calculé à partir de la valeur prise à la moitié de la durée vocalique * 4 répétitions* 4 contextes confondus* 20 locutrices	214
Figure 99 : Interface graphique dans Praat du test de perception interlangue des voyelles isolées.....	219
Figure 100 : Triangles vocaliques F1/F2, F2/F3 (en Bark) des voyelles orales du français produites en syllabe finale du logatome kVkvkVkvk (4 répétitions) par dix Françaises (trait plein) et dix Tchèques (trait pointillé). Les ellipses de dispersion sont tracées à un écart type. Notons que le [ø] des Tchèques se trouve à plus d'un écart type de la « référence ».....	244
Figure 101 : Valeurs moyennes de F1, F2, F3 et F4 (en Bark) du [i] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un losange) et relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1	248
Figure 102 : Valeurs moyennes de F1, F2, F3 (en Bark) du [ε] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un losange) et relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1	250

Figure 103 : Valeurs moyennes de F1, F2 (en Bark) du [u] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un losange) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1.....	251
Figure 104 : Valeurs moyennes de F1, F2, F3 (en Bark) du [e] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions de dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un losange) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 2.....	254
Figure 105 : Valeurs moyennes de F1, F2 du [o] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un triangle) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1.....	255
Figure 106 : Valeurs moyennes de F1, F2, F3 (en Bark) de [a] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions de dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un triangle) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1.....	257
Figure 107 : Valeurs moyennes de F1 et F2 (en Bark) de [ɔ] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions de dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un triangle) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 2.....	260
Figure 108 : Valeurs moyennes de F1, F2 et F3 du [œ] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un triangle) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 2.....	262
Figure 109 : Valeurs moyennes de F1, F2 et F3 du [y] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un triangle) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1.....	265
Figure 110 : Valeurs moyennes de F1, F2 et F3 du [ø] en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 4 répétitions par dix Françaises (marquées par une croix) et 10 Tchèques (marquées par un triangle) et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1.....	266
Figure 111 : Triangles vocaliques F1/F2, F2/F3 (en Bark) des voyelles orales du français produites en syllabe finale des logatomes RVRVRVR par dix Françaises (marquées par un rond) et dix Tchèques (marquées par un triangle), 4 répétitions. Les ellipses de dispersion sont tracées à un écart type.....	268
Figure 112 : Triangle F1/F2 (en Bark) des voyelles orales du Français produites par 10 Tchèques en isolation (marquées par un rond), en contextes labial (triangle), coronal (croix), palato-vélaire (losange) et uvulaire (triangle vers le bas). Moyennes calculées à partir de 4 répétitions* syllabe finale* valeur prise à la moitié de la durée vocalique / 4 répétitions (logatomes) * 3 valeurs : à 1/3, ½ et 2/3 de la durée vocalique pour les voyelles isolées. Remarquons une grande proximité acoustique des voyelles moyennes insérées dans le même contexte phonétique	274
Figure 113 : F1 et F2 des 120 sons vocaliques synthétisées avec Klatt (carrés rouge). Les voyelles [i, a, u] y servent de point de repère	289
Figure 114 : Interface graphique du test d'identification des sons vocaliques synthétisées avec Klatt	290
Figure 115 : Sons vocaliques synthétisées avec Klatt catégorisés comme /e, ε, ø, œ, o, ɔ/ avec un <i>fit index</i> de 4 et plus sur 5. Ceux catégorisés avec un <i>fit index</i> de 5 sont marqués en grand. Les voyelles [i, a, u] servent de point de repère	294
Figure 116 : Questionnaire phonologique portant sur la façon de prononcer les deux mots de chaque paire minimale dont le contraste repose sur les voyelles moyennes	296
Figure 117 : Contraste acoustique entre les voyelles e/ε produites en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ]. Valeurs de F1, F2 et F3 (en Bark) sont calculées à partir de 4 répétitions par 10 Françaises natives et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1.....	299

Figure 118 : Spectrogramme des voyelles [e] et [ɛ] produites par une Française native en isolation et en contextes symétriques p, t, k, R (syllabe finale).....	300
Figure 119 : Contraste acoustique entre les voyelles o/ɔ produites en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ]. Valeurs de F1, F2 (en Bark) sont calculées à partir de 4 répétitions par 10 Françaises natives et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1301	
Figure 120 : Spectrogramme des voyelles [o] et [ɔ] produites par une Française native en isolation et en contextes symétriques p, t, k, R (syllabe finale).....	303
Figure 121 : Contraste acoustique entre les voyelles ø/œ produites en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ]. Valeurs de F1, F2 et F3 (en Bark) sont calculées à partir de 4 répétitions par 10 Françaises natives et relevées à un tiers, à la moitié et 2 tiers de la voyelle. L'écart type tracé est de 1.....	304
Figure 122 : Spectrogramme des voyelles [ø] et [œ] produites par une Française native en isolation et en contextes symétriques p, t, k, R (syllabe finale).....	305
Figure 123 : Triangle F1/F2 (1 ^e image) et F2/F3 (2 ^e image) des voyelles isolées du français (en Bark) produites par 10 Tchèques (trait pointillé, 10 répétitions) et 10 Françaises natives (trait plein, 4 répétitions). Les voyelles [ø] et [ɛ] des Tchèques se trouvent à moins d'un écart type de la "référence"	313
Figure 124 : Spectrogramme du [ø] isolé produit par les locutrices tchèques 1, 2, 4, 5, 6, 7, 8, 9 et 10, majoritairement identifié par les Français natifs comme tel (avec un <i>fit index</i> de 4 et plus sur 5).....	314
Figure 125 : Spectrogramme du [ø] isolé produit 10 fois par la locutrice tchèque T3	315
Figure 126 : Triangles F1/F2 et F2/F3 (en bas) des voyelles isolées du français (en Bark) prononcées 10 fois par la locutrice T3 (en trait pointillé) et 4 fois par 10 Françaises natives non-méridionales (en trait plein). L'écart type affiché est de un	316
Figure 127 : Spectrogramme du [ɛ] isolé produits par les locutrices tchèques T2, T3, T4, T5, T6, T9 et T10, majoritairement identifié par les Français natifs comme tel (avec un <i>fit index</i> de 4 et plus sur 5).....	317
Figure 128 : Spectrogramme du [e] isolé produit 10 fois par la locutrice tchèque T9.....	319
Figure 129 : Triangles F1/F2 et F2/F3 des voyelles isolées du français (en Bark) prononcées 10 fois par la locutrice T9 (en trait pointillé) et 4 fois par 10 Françaises natives non-méridionales (en trait plein). L'écart type affiché est de un	320
Figure 130 : Spectrogramme de dix répétitions du [o] isolé produit par la locutrice tchèque T9.....	321
Figure 131 : Triangle F1/F2 (en Bark) des voyelles isolées du français prononcées 10 fois par la locutrice T9. Remarquons la stabilité dans la production du timbre vocalique [o].....	322
Figure 132 : Spectrogramme de dix répétitions de [œ] isolé produites par la locutrice tchèque T3.....	323
Figure 133 : Triangles F1/F2 et F2/F3 (en bas) des voyelles isolées du français (en Bark) prononcées 10 fois par la locutrice T3 (en trait pointillé) et 4 fois par 10 Françaises natives non-méridionales (en trait plein). L'écart type affiché est de deux	325
Figure 134 : Spectrogramme de dix répétitions de [œ] isolé produit par la locutrice tchèque T4	325
Figure 135 : Triangles F1/F2 et F2/F3 (en bas) des voyelles isolées du français (en Bark) prononcées 10 fois par la locutrice 4	327
Figure 136 : Spectrogramme de dix répétitions de [ɔ] isolé produit par la locutrice tchèque T9	328
Figure 137 : Triangle F1/F2 (en Bark) des voyelles isolées du français prononcées 10 fois par la locutrice T9	329
Figure 138 : Valeurs moyennes de F1, F2 et F3 (en Bark) de e/ɛ produit par la locutrice T8 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. Il est à noter la différence de valeurs de F1 et F2 en syllabe finale tVt.....	331
Figure 139 : Triangle F1/F2 (en Bark) des voyelles du français prononcées 10 fois par la locutrice T8 en syllabe finale tVt (valeur prise à la moitié de la durée vocalique). Il est à noter que e /ɛ sont pour la plupart séparés sur le plan F1 et F2	332
Figure 140 : Valeurs moyennes de F1, F2 et F3 (en Bark) de e/ɛ produit par la locutrice T9 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. Notons la différence de valeurs de F1 en isolation et en syllabe finale pVp, tVt et kVk	332

- Figure 141 : Valeurs moyennes de F1, F2 et F3 (en Bark) de e/ε produit par la locutrice T10 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. L'écart type tracé est de 1. Notons la différence de valeurs de F1 et/ou F2 en isolation et en syllabe finale pVp, kVp et RVR..... 333**
- Figure 142 : Valeurs moyennes de F1, F2 et F3 (en Bark) de e/ε produit par la locutrice T2 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. L'écart type tracé est de 1. Notons la différence de valeurs de F2 en syllabe finale pVp et tVt..... 334**
- Figure 143 : Triangle F1/F2 (en Bark) des voyelles du français prononcées 10 fois par la locutrice T2 en syllabe finale pVp (valeur prise à la moitié de la durée vocalique). Notons que les occurrences des e /ε sont pour la plupart séparés sur le plan F2 334**
- Figure 144 : Triangle F1/F2 (en Bark) des voyelles du français prononcées 10 fois par la locutrice T2 en syllabe finale tVt (valeur prise à la moitié de la durée vocalique). Notons que e (en vert)/ ε (en bleu) sont séparés sur le plan F2 335**
- Figure 145 : Valeurs moyennes de F1, F2 et F3 (en Bark) de ø/œ produit par la locutrice T8 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. L'écart type tracé est de 1. Notons la différence de valeurs de F1 et/ou F2 en isolation et en syllabe finale pVp et tVt 336**
- Figure 146 : Valeurs moyennes de F1, F2 et F3 (en Bark) de ø/œ produit par la locutrice T9 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. L'écart type tracé est de 1. Notons la différence de valeurs de F1 et/ou F2 en isolation et en syllabe finale tVt..... 336**
- Figure 147 : Valeurs moyennes de F1, F2 et F3 (en Bark) de o/ɔ produit par la locutrice T10 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. L'écart type tracé est de 1. Notons la différence de valeurs de F1 et/ou F2 des voyelles isolées et dans des syllabes finales des quatre contextes 338**
- Figure 148 : Valeurs moyennes de F1, F2 et F3 (en Bark) de o/ɔ produit par la locutrice T1 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. L'écart type tracé est de 1. Notons la différence de valeurs de F1 en syllabe finale kVp 338**
- Figure 149 : Valeurs moyennes de F1, F2 et F3 (en Bark) de o/ɔ produit par la locutrice T2 en contexte nul (traits horizontaux) et dans des logatomes CVCVCVC où C = [p, t, k, ʁ], calculées à partir de 10 répétitions. Valeurs relevées à un tiers, à la moitié et deux tiers de la voyelle. L'écart type tracé est de 1. L'écart type tracé est de 1. Notons la différence de valeurs moyennes de F1 en syllabe finale kVp..... 339**
- Figure 150 : Triangle vocalique des voyelles orales isolées du français (en pointillé) et du tchèque (en trait plein) sur le plan F1/F2 (en Bark). Les ellipses de dispersion sont tracées à 2 écarts types de la moyenne calculée à partir de productions de 10 Françaises/ 20 Tchèques* 4 répétitions* 3 valeurs par voyelle..... 343**
- Figure 151 : Valeurs moyennes F1/F2 (en Bark) des voyelles orales du français produites dans des logatomes CVCVCVC où C = /p, t, k/, calculées à partir de 4 répétitions par dix Françaises (valeurs prises à la moitié vocalique)..... 346**
- Figure 152 : Valeurs moyennes de F1/F2 (en Bark) des voyelles orales du français produites dans des logatomes CVCVCVC (C = p, t, k, R ; tous contextes confondus) par dix futures enseignantes (marquées par un triangle) et des voyelles du tchèque produites dans des logatomes CVCVCVC (C = p, t, k, h ; tous contextes confondus) par 20 Tchèques (marquées par un rond), 4 répétitions. Les ellipses de dispersion sont tracées à un écart-type 347**

Perception et production des voyelles orales du français par des futures enseignantes tchèques de Français Langue Étrangère (FLE)

Résumé

Cette étude acoustico-perceptive concerne les limites de la perception et de la production des voyelles orales du français [i, e, ε, a, u, o, ɔ, y, ø, œ], en isolation et en contextes consonantiques divers, chez dix tchécophones, futures enseignantes de Français Langue Étrangère (FLE). Les résultats montrent que (1) La maîtrise phonétique des voyelles dépend de leurs graphies et de l'entourage consonantique. (2) Les voyelles fermées [i, y, u] et le [a] sont globalement maîtrisées avec authenticité. (3) Les capacités de perception des contrastes entre les voyelles moyennes e/ε, ø/œ et o/ɔ ainsi que leur production sont limitées. Ces résultats ne sont que partiellement en accord avec les prédictions établies à partir du Speech Learning Model (SLM) de Flege (1995), basé sur la notion de *similarité* phonétique qui existe entre la langue maternelle (LM) et la langue étrangère (LE).

Mots clés : *phonétique, acoustique, formants, coarticulation, perception, prononciation, voyelles, français, tchèque, Français Langue Étrangère (FLE), Speech Learning Model (SLM)*

Perception and Production of French Oral Vowels in Pre-Service Czech Teachers of French as a Foreign Language (FFL)

Abstract

This acoustic-perceptual study concerns the limits of perception and production of French oral vowels [i, e, ε, a, u, o, ɔ, y, ø, œ], in isolation and in different consonantal contexts, in ten pre-service Czech teachers of French as a Foreign Language (FFL). The results show that (1) Phonetic proficiency in vowels depends on their spellings and consonantal context. (2) Vowels [i, y, u] and [a] are generally mastered with authenticity. (3) The ability to hear contrasts between the vowels e/ε, ø/œ and o/ɔ and pronounce them is limited. These results are only partially consistent with the predictions established in the Speech Learning Model (SLM) by Flege (1995), based on the notion of phonetic *similarity* between the mother tongue (MT) and the foreign language (FL).

Keywords : *phonetics, acoustics, formants, coarticulation, perception, pronunciation, vowels, French, Czech, French as a Foreign Language (FFL), Speech Learning Model (SLM)*

UNIVERSITÉ SORBONNE NOUVELLE – PARIS 3

École Doctorale 268 « Langage et Langues : Description, théorisation, transmission »

UFR de Linguistique et Phonétique Générales et Appliquées (ILPGA)

1 rue Censier, 75005 Paris